GROUP THEORY OR NO GROUP THEORY: UNDERSTANDING SELECTION RULES IN ATOMIC SPECTROSCOPY

by

Joshua Robert Hunt

B.S. in Chemical Engineering, University of Pittsburgh, 2014

BPhil in Philosophy, University of Pittsburgh, 2014

Submitted to the Graduate Faculty of

The Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

BPhil in Philosophy

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

THE DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Joshua Robert Hunt

It was defended on

March 27, 2014

and approved by

Robert Batterman, Professor, Philosophy

Mark Wilson, Professor, Philosophy

Laura Ruetsche, Professor, Philosophy

Thesis Director: Kenneth Manders, Associate Professor, Philosophy

GROUP THEORY OR NO GROUP THEORY: UNDERSTANDING SELECTION RULES IN ATOMIC SPECTROSCOPY

Joshua Robert Hunt, BPhil

University of Pittsburgh, 2014

In the late 1920's and early 1930's, physicists applied group representation theory to the quantum mechanics of atomic spectra. At the same time, physicists developed an alternative approach to theoretical atomic spectra that avoids using group theory. These two approaches exhibit nontrivial intellectual differences: the group theoretic approach provides a deeper understanding of many phenomena in atomic spectra. By focusing on derivations of selection rules for atomic spectra, I explicate one case where group theory enhances understanding. I refer to the non-group theoretic approach as the commutator approach; it serves as a benchmark for comparison. This case study motivates a deflationary account of mathematical explanations in science. According to my account, both group theoretic and commutator derivations explain selection rules for atomic spectra. I use these derivations to problematize stronger accounts of mathematical explanation that rely on a notion of relevance. Arguing that selection rules are an example of universality, I also criticize a strong interpretation of Batterman and Rice's minimal model account of explanations of universality.

After examining these accounts of explanation, I argue that explanatory criteria do not distinguish the intellectual content of the group theoretic and commutator approaches. Instead, I develop an account of scientific understanding that distinguishes these approaches based on organizational differences. Adopting terminology from Manders, I argue that these organizational differences arise from differences in the approach's expressive means. Group theory reorganizes selection rule derivations by re-expressing physical concepts more effectively than the commutator approach. I argue that this superior organizational structure accounts for how group theory provides a heightened understanding of selection rules.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

This thesis taught me that the quickest way up a mountain is not necessarily the path of steepest descent. After sliding down that path a few times, I finally gained enough sense to take a different tack. I meandered around that mountain valley for some time before climbing to a comfortable precipice. I offer a view neither from the top nor the bottom, but somewhere in the middle (I would like to think the lower middle). Along the way, I received the help, advice, and support of a number of generous individuals. It is a pleasure to thank them.

First and foremost, I thank Professor Ken Manders for his overflowing support, generous training, and continued guidance. This thesis began as a term paper for his Philosophy of Mathematics class. Working under him has been a terrific intellectual experience and a great honor. My intellectual debt to him will be lifelong.

I am extremely grateful for the assistance of the members of my thesis defense committee. Bob Batterman took excellent care of me in the final stages of this project. His comments forced me to significantly restructure my argument. I owe him special thanks. I also thank Mark Wilson and Laura Ruetsche, especially for their critical comments and questions during my defense. I hope to one day be half as much a philosopher as they are.

Along with Professor Manders, Rebecca Morris patiently listened to me explain the details of my project during an extended case study seminar. I have benefitted enormously from her criticisms, comments, and probing questions. Her own work has served as a great inspiration.

For financial support, I thank the University of Pittsburgh Honors College for sponsoring Brackenridge Research Fellowships in the summers of 2012 and 2013 and an undergraduate research fellowship in the Spring of 2013. I especially thank Dean Edward Stricker, Dr. Peter Koehler, Karen Billingsley, Dave Hornyak, Nate Hilberg, and Mike Giazzoni. I thank members of the Brackenridge Research communities for their stimulating questions and continued criticism. I was pleasantly surprised—and mildly alarmed—to learn that some peers thought I was engaged in a science project.

Parts of this thesis were presented at the fourth biennial Society for Philosophy of Science in Practice conference during a symposium on Mathematical (and applied mathematical) Conceptual Practice. Members of the audience provided helpful comments. I particularly thank Bihui Li, Chris Pincock, Kareem Khalifa, Joseph, Berkovitz, James Brown, David Stump, and Andrea Woody.

For their intellectual support throughout my research, I thank Julia Bursten, Giovanni Valente, and Professor Thomas Ricketts. I am honored to have been Julia's "undergrad;" I have learned a lot from her keen editorial eye. I thank Professor Andrew Daley for generous help in the early stages of this project. Professors Robert Devaty, Brian D'Urso, Chandralekha Singh, Sandy Asher, Geoffrey Hutchison, and Dr. Eugene Wagner all assisted in some way or another. I especially thank Professor Edward Gerjuoy for sitting down to tell me about the good old days of atomic spectroscopy. I thank Professor Katherine Brading for her comments on an earlier draft.

For a varied mix of emotional and intellectual support, I thank Matthew Schaff, Simon Brown, Evan Arnet, Carsen Stringer, Rohith Palli, Dave Korotky, Yara Skaf, Tommy McFarlin, Robert Hunt, Roberta Hunt, Thomas Hunt, Colleen Hunt, Susan Carroll, Sean Hunt, and Emily Hunt. These latter seven deserve a special thanks their lifelong—and mystifyingly unconditional—support of my intellectual endeavors. For moral support throughout the duration of this project—and I mean duration—I thank Jessy Volkert. This thesis could not have been completed without the support and flexibility of my favorite chemical engineers: thank you Gerald McFarlin, Joseph Andros, and Robert "The Man" Briggs. Members of the University of Pittsburgh undergraduate philosophy of science club have had the forbearance to put up with me for three years. With or without consent, they have discussed numerous parts of this thesis.

Finally, I thank Arvind Prasadan and Florica Constantine, for everything.

1.0 GROUP THEORY OR NO GROUP THEORY

1.1 INTRODUCTION

The mathematical theory of groups and their representations has found widespread application in science. Among diverse fields, group theory has proven fruitful in crystallography, quantum mechanics, and particle physics (Cornwell 1984). It is often said that the applicability of group theory stems from its treatment of physical symmetry: group theory provides a natural language for reasoning about symmetries. A great deal more can be said. When group theory is fruitfully applied to a physical problem, there is a shared sense that an act of tremendous intellectual power has taken place. These applications represent a significant advance in our knowledge and understanding of physical phenomena. This thesis considers in detail one case where group theory affects an intellectual advance. Although examples can be multiplied, we gain much by looking closely at how intellectual progress is wrought. In this case, there is an alternative mathematical approach that avoids using group theory. This alternative approach facilitates an articulation of what group theory contributes intellectually.

One of the earliest applications of group theory was to a theoretical account of atomic spectroscopy. Physicists developed this theory in the late 1920s and early 1930s as a testing ground for the nascent quantum mechanics. Eugene Wigner (1927) and Hermann Weyl (1927) published the first papers detailing the application of group theory to quantum mechanics. Weyl published a monograph in 1928, which was revised and expanded by 1931. Wigner and John von Neumann collaborated on a series of papers that further developed a group theoretic approach (1928a, b, c). Wigner compiled this work into an introductory text in 1931. Bartel van der Waerden (1932) also published a monograph. These works—which are still read fruitfully today—demonstrate how quickly applied group theory reached maturity in theoretical atomic spectroscopy. Its application spurred numerous other developments in molecular and nuclear

physics in the 1930s and 40s. As such, it is surprising that physicists in general were slow to adopt group theoretic methods in atomic spectroscopy. In the theory of atomic spectra, physicists viewed group theory as largely unnecessary for more than 20 years after its initial application.

This widespread dismissal of group theory was motivated largely by the success of an alternative approach to atomic spectra that made no explicit use of groups. In their monumental 1935 text (reprinted in 1964), *The Theory of Atomic Spectra*, E. U. Condon and G. H. Shortley gave a successful account of atomic spectra that remained the standard textbook on the subject for a few decades. B. R. Judd, a popularizer of the group theoretic approach, has called Condon and Shortley's text "a volume seemingly insusceptible of improvement" (Judd 1963, v). In particular, their third chapter relies heavily on algebraic relationships between quantum mechanical operators. These commutation relations stem from the quantum theory of angular momentum, introduced by Born, Heisenberg, and Jordan (1926) and further developed by Born and Jordan (1930) and Güttinger and Pauli (1931).

These two approaches—the group theoretic and non-group theoretic—afford two ways of understanding numerous problems in atomic spectra. Unlike many cases of alternative physical theories, these approaches do not compete with each other. It is not the case that only one of them provides a correct theory of atomic spectra. Although the early years of quantum mechanics witnessed methodological tension between the two approaches, they were never held as incompatible theories. In their text, Condon and Shortley acknowledge the work done by Wigner, Weyl, and van der Waerden, directing the interested reader to their monographs. Mathematically, these two approaches are related aspects of Lie theory. The group theoretic approach uses mathematical structures known as Lie groups—topological groups that support the methods of calculus. The non-group theoretic approach avoids discussion of groups by implicitly utilizing Lie algebras. These mathematical structures are vector spaces equipped with a commutator product. Each Lie group gives rise to a Lie algebra. In special cases, much of the information about a Lie group is recoverable from its Lie algebra (Cornwell 1984).

Despite this mathematical connection between the two approaches, their treatments of atomic spectra are distinctive. In the end, they recover many of the same phenomena, but the paths they take are markedly different. They employ alternative mathematical concepts and solution procedures, giving rise to distinct characters of thought. To make these differences amenable to philosophical analysis, I focus on a specific class of phenomena: selection rules in

atomic spectra. Selection rules provide restrictions governing how atoms transition from one atomic state to another. During these transitions, an electron undergoes a change in energy level, and radiation is either absorbed by or emitted from an atom. Although only a small piece of the grand edifice that is our theory of atomic spectra, selection rules are of immense empirical and theoretical importance.[1] Empirically, selection rules are necessary for the proper interpretation of the information encoded in spectra (Engel and Hehre 2010, p. 218). Spectroscopy itself performs a crucial epistemic function as the source of "most of the direct experimental information we have about the structure of atoms and molecules" (Flurry 1980, p. 160).

Methodologically, selection rules provide an attractive case study because physicists have thoroughly treated them from both the group theoretic and non-group theoretic vantage points. On the non-group theoretic side, early presentations are found in Born, Heisenberg, and Jordan (1926), Dirac (1930) and subsequent editions, and Condon and Shortley (1935). This viewpoint has also survived in more modern treatments (Griffiths 2005). Since these non-group theoretic treatments proceed primarily through commutation relations, I refer to them as the commutator approach to selection rules. On the group theoretic side, Wigner (1927, 1931) and van der Waerden (1932) include treatments. Most modern expositions of applied group theory in quantum mechanics derive selection rules as an illustrative example of group theoretic methods (Tinkham 1964; Petrashen and Trifonov 1969).

In Chapter 2, I present derivations of selection rules from both the commutator and group theoretic viewpoints. These derivations aim to answer the following why-question: why do atoms of low atomic number exhibit the same selection rules upon excitation? This question is an instance of what Batterman (2002) calls a type (ii) explanatory why-question: "A type (ii) why-question asks why, in general, patterns of a given type can be expected to obtain" (p. 23). Here, we are interested in why atoms of different atomic number exhibit the same pattern of selection rules. In their own ways, both the commutator and group theoretic approaches provide answers to this question.

[1] For an illuminating historical study of the development of selection rules and Wigner's early application of group theory, see Borrelli (2009).

1.2 DISTINGUISHING EXPLANATION AND UNDERSTANDING

The lack of a unique method for deriving selection rules facilitates an analysis of mathematical explanations in science and scientific understanding. Since the commutator and group theoretic derivations do not compete with each other, we can accept them both as legitimate methods for deriving selection rules. I refer to this lack of competition as compatibility: the two approaches to deriving selection rules are compatible. If the two approaches were incompatible, we could distinguish them based on which one—if either—we should ultimately accept. Thanks to compatibility, we do not need to exclusively choose one approach over the other when it comes to explaining and understanding selection rules. Compatibility raises more subtle questions about how the approaches differ regarding explanation and understanding. In particular, I examine whether or not either approach should be viewed as explanatory; ultimately, I argue that both approaches provide explanations of selection rules. Although both approaches explain selection rules, they exhibit a decisive intellectual asymmetry. Introducing group theoretical concepts enables a more illuminating treatment of selection rule phenomena, providing enhanced understanding. In articulating the source of this enhanced insight, the commutator approach serves as a useful benchmark for comparison.

In Chapters 3 and 4, I distinguish explanation from understanding. My terminological distinction is motivated by the task of articulating how group theory provides enhanced understanding. Prior to analysis, it is conceivable that the group theoretic approach is genuinely explanatory of selection rules while the commutator approach is explanatorily deficient. If true, this account would provide a simple story about how group theory enhances understanding: the group theoretic account alone would provide explanations of selection rules. I resist this way of interpreting the derivations. In Chapter 3, I argue that both approaches should be seen as explanatory. In doing so, I aim to separate our philosophical term "genuine explanation" into two components: "explanation" and "understanding." I defend a minimal account of explanation according to which an argument is explanatory if it recovers a phenomenon of interest according to principled constraints. Thus, when I refer to an argument as explanatory, I do not have in mind a richer sense of genuine explanation. I believe that much of the richness present in "genuine explanation" is better dealt with under the term "scientific understanding." My account of

understanding characterizes insight in terms of organizational virtues such as modularization, tractability, and uniformity of treatment.

Although my use of the words "explanation" and "understanding" are nonstandard, I believe that this way of speaking facilitates an analysis of how mathematics provides insight in science. When philosophers debate which of two compatible approaches is genuinely explanatory of a phenomenon, it is more fruitful to take both approaches as explanatory and analyze their intellectual differences. These differences pertain to differences in re-expressing physical problems, restructuring solution procedures, and reorganizing relationships between physical and mathematical constraints. Although they account for much of the richness of our usual notion of explanation, they have little to do with "explanation" in the weaker sense that I employ. Hence, I find it convenient to separate these two notions, even though they are usually intertwined in discussions of genuine explanation.

In Chapter 3, I motivate my minimal restriction on explanation through an analysis of more restrictive accounts of explanation. I consider a family of positions that I refer to as relevance accounts of explanation. Relevance accounts contend that an argument is explanatory only when it references relevant physical and mathematical features while eliminating irrelevant details. If relevance accounts are correct, then we should determine which of the commutator or group theoretic approaches better satisfies these restrictions on explanation. For the sake of argument, assume that upon investigation we would find that the group theoretic approach eliminates more irrelevant details and references a greater number of relevant features. Relevance accounts would legitimize the group theoretic approach as providing a more genuine explanation. Furthermore, relevance accounts would locate the superior understanding provided by the group theoretic approach in the elimination of irrelevant details and greater reliance on relevant features.

In Section 3.3.1, I resist this characterization by problematizing relevance accounts. Relevance accounts contend that determinations of relevance ground explanations. I argue that—at least in the context of my case study—it is my minimal sense of explanation that grounds determinations of relevance. This introduces a circularity problem for relevance accounts: relevance should not be taken to ground explanation if explanatory success grounds relevance. Section 3.3.2 develops a related problem arising from the existence of multiple compatible explanations of a given phenomenon. Since compatible explanations rely on different sets of

physical and mathematical features, relevance accounts seem committed to determining a unique set of relevant features. More charitably, relevance accounts must determine which of two sets of mathematical and physical features references a greater number of relevant details and eliminates a greater number of irrelevant details. To make these determinations of relevance, we would rely on my minimal notion of explanation. This strengthens the circularity problem introduced in Section 3.3.1.

In the specific context of explanations of universality, Batterman and Rice (unpublished) develop an alternative account of explanation. In physics, universality refers to a pattern of behavior exhibited by a class of physical systems whose members are constitutionally distinct. For instance, both water and gasoline—although micro-structurally distinct—exhibit a parabolic momentum profile under laminar flow conditions. A parabolic momentum profile is a universal behavior exhibited by a universality class of fluids, including water and gasoline. Explanatory questions about universality are a particular kind of type (ii) why-question: they ask why a given universal behavior obtains in this class of systems. Section 3.4 presents the details of Batterman and Rice's account, which focuses specifically on explanations of universality. I argue that selection rules are properly seen as a kind of universal behavior, making Batterman and Rice's account germane to my analysis of explanation in the context of selection rules. Batterman and Rice criticize relevance accounts—which they refer to as "common features accounts"—for a similar reason to the circularity challenge that I develop.

However, Batterman and Rice place more restrictions on explanations of universality than the deflationary account of explanation that I defend. In Section 3.5, I consider two interpretations of Batterman and Rice's prescriptive claims regarding genuine explanations of universality. On a literal interpretation of Batterman and Rice's restrictions, explanations of universality must undertake a stability analysis that identifies a "minimal model" and delimits its universality class. This literal interpretation is in tension with scientific practice. Most, if not all, published derivations of selection rules do not conduct the kind of stability analysis prescribed by Batterman and Rice. To defuse this tension, I propose a weaker interpretation of Batterman and Rice's account of explanation. I argue that we should interpret Batterman and Rice as offering a methodological prescription for good explanatory practice, rather than a strict requirement for genuine explanation. This prevents my minimal account of explanation from conflicting with Batterman and Rice's stricter account in the context of selection rules. Even

though derivations of selection rules do not satisfy Batterman and Rice's requirements, we can view them as explanatory nonetheless. At the same time, scientists could potentially provide further justification for the usual explanations of selection rules by meeting the criteria laid out by Batterman and Rice.

Having defended both the commutator and group theoretic approaches as explanatory, I turn in Chapter 4 to the question of how they differ in the understanding they provide. I begin in Section 4.1 by disentangling two senses in which an explanation can be better than an alternative explanation. Assuming that both explanations are empirically adequate, we might compare their relative theoretical virtues, such as simplicity, elegance, and fruitfulness. Appeal to theoretical virtues is a standard move when using inference to the best explanation (IBE) to allay underdetermination problems. Some scientific realists take these theoretical virtues to be truth-tracking, thereby providing an epistemic means for preferring an explanation or theory among competing alternatives. This strategy relies on a first sense of "better explanation," according to which an explanation is better if it is closer to the truth. This first sense of "better explanation" is not relevant for characterizing the intellectual features that arise in my case study. Unlike competing alternatives, derivations of selection rules from the commutator and group theoretic approaches are compatible. Hence, we do not need to determine which approach is closer to the truth. Instead, we should focus on a second sense of "better explanation" that emphasizes how compatible explanations can differ in the insight they provide. In this context, theoretical virtues remain important, but their importance is divorced of any purported relationship to truth. When I return to these virtues in Section 4.3, I argue that they are organizational in nature: they deal with how an approach structures the solution to a problem.

The commutator and group theoretic approaches differ in their organizational virtues because they rely on different conceptual resources to explain selection rules. Section 4.2 undertakes a characterization of these conceptual differences and their consequences for the phenomena that both approaches can address. Adopting terminology introduced by Manders (unpublished), I use the phrase "expressive means" to denote the mathematical and physical concepts employed by an approach. Because alternative concepts sometimes make accessible the same phenomenon, I distinguish expressive means from expressive power. I use "expressive power" to denote the set of phenomena that a conceptual framework can talk about. Differences in expressive power arise from differences in expressive means. The group theoretic approach is

able to characterize symmetries of atomic systems that remain inaccessible to the commutator approach. This difference in expressive power enables the group theoretic approach to restructure the problem of deriving selection rules. In Section 4.3, I argue that group theory deepens our understanding of selection rules by providing a more effective organizational structure for reasoning about them.

2.0 TWO APPROACHES TO SELECTION RULES

In Chapter 1, I introduced the group theoretic and commutator approaches as compatible frameworks for deriving selection rules. In Chapters 3 and 4, I use these compatible approaches to advance an account of scientific explanation and understanding. These analyses are supported by the nitty-gritty details of the commutator and group theoretic derivations. Hence, before proceeding, I collect relevant background information for understanding these derivations. Section 2.1 introduces selection rules in atomic spectra for hydrogenic and multi-electron atoms. Readers who are comfortable with the basic theory of atomic spectra can safely skim or skip most of this section, with the exception of 2.1.5, where I characterize selection rules as an example of universality. Section 2.2 details the commutator approach to selection rules in atomic spectroscopy. It is not necessary to follow every step in the derivations. They are provided for completeness, and through them I aim to convey a sense of what goes on in this approach. Section 2.3 presents an alternative group theoretical approach to selection rules. It would be natural to proceed further and consider group theoretic and commutator approaches to the Wigner-Eckart theorem. This theorem extends further our understanding of selection rules. Nonetheless, the remarks made in Chapters 3 and 4 do not require these additional details, so I leave them aside for future work.

2.1 SELECTION RULES IN ATOMIC SPECTROSCOPY

In the most general sense, spectroscopy consists in the recorded interaction of electromagnetic radiation with matter. Upon excitation—e.g. by means of heating—chemical substances such as atoms or molecules emit radiation of discrete wavelengths. Different chemical substances yield

characteristic emission spectra,[2] enabling the composition of stars, nebulae, and other light-emitting bodies to be identified.

The emitted wavelengths of electromagnetic radiation are commonly expressed via wave numbers, υ, which are the reciprocals of the corresponding wavelengths: $\upsilon = 1/\lambda$, commonly expressed in units of cm^{-1}. In a typical atomic spectrum, such as that of a hydrogen atom, the lines occur in collections called "series." The lines within a given series exhibit a regular decrease in their spacing and eventually converge to a limit. The first series of lines discovered for hydrogen lies in the visible range (consisting of wavelengths from around 385 to 780 nm) and is known as the Balmer series. Balmer discovered that the wave numbers can be characterized by the formula

$$v = R\left(\frac{1}{2^2} - \frac{1}{n_1^2}\right)$$

where R is the Rydberg constant—with value of 109,737.3 cm^{-1}—and n_1 is a natural number taking values 3, 4, 5, etc.

The Balmer series is the result of atomic transitions from excited states of hydrogen to the second lowest energy state, characterized by $n_2 = 2$. Similarly, the Pfund series occurs when the electron orbiting hydrogen falls to the fifth lowest energy level: n = 5. The wavenumbers of its spectral lines are given by the analogous formula $v = R\left(\frac{1}{5^2} - \frac{1}{n_1^2}\right)$ where n_1 now ranges over 6, 7, 8, etc.

Generalizing these formulae, spectroscopists began to write wave numbers as the difference between two spectroscopic terms: $v = T_2 - T_1$, where the terms generally have more complicated formulas than those for hydrogen. The empirical Rydberg-Ritz combination principle states that the difference between two spectroscopic terms yields a wave number corresponding to a spectral line. If this principle held exactly, then many more spectral lines would be observed than are in practice. In fact, only wave numbers corresponding to certain combinations of terms occur with appreciable intensity. This means that transitions occur only between certain states of the atom: given an initial state, there are restrictions on the final state after emission of radiation. These restrictions are characterized by the selection rules for spectral

[2] Depending on the method of excitation, the emission spectra of a single substance will exhibit variations in the intensity of the lines, although these variations do not normally lead to ambiguity in identifying the substance.

transitions. While most of the selection rules admit numerous exceptions and have restricted domains of validity, they have proved extremely useful in the interpretation and characterization of spectra.

2.1.1 Stationary States for Hydrogenic Atoms

Selection rules are formulated using indices known as quantum numbers, which characterize stationary states of an atom. Mathematically, quantum numbers are the eigenvalues of operators which commute with the system's Hamiltonian—an operator that describes the energy of the system. In virtue of commuting with the Hamiltonian, these operators represent observable quantities that are constants of motion, i.e. quantities which are conserved over time. Often, the eigenvalues of a given observable are degenerate. This means that multiple stationary states possess the same value for this observable. However, these stationary states differ in general in at least one eigenvalue, i.e. they are described by a different set of quantum numbers. By finding a complete set of commuting observables, it is possible to label all of the possible energy states of a system.

For hydrogen and other hydrogenic (one-electron) atoms, the electron moves in a coulomb potential. Coulomb potentials are spherically symmetric, and neglecting relativistic and spin effects, the energy of the electron depends only on the radial distance from the atomic nucleus. The Hamiltonian operator corresponds to the energy of the atomic state, and the eigenvalue of the Hamiltonian is denoted by *n*, referred to as the principal quantum number. The principal quantum number takes integer values 1, 2, 3, etc., a manifestation of the fact that the energy of the atom is quantized so long as the electron remains bound.

When relativistic and spin effects are neglected, the orbital angular momentum operator, L^2, commutes with the Hamiltonian. This reflects the rotational symmetry of the electron in the coulomb potential, which gives rise to the conservation of angular momentum. Formally, the eigenvalues of the orbital angular momentum operator are given by $\hbar[l(l+1)]^{1/2}$, where $\hbar$ equals Planck's constant divided by 2π, and l is an integer which takes values from 0, 1, 2, . . . , n-1. These eigenvalues represent the magnitude of the orbital angular momentum of the electron. Since $[l(l+1)]^{1/2} \approx l$, the orbital angular momentum is often approximated by $l\hbar$, and l is used as the quantum number, commonly referred to as the azimuthal quantum number.

The angular momentum is a vector quantity with components in three directions. However, as a consequence of the Heisenberg uncertainty principle, only one of these three directional components can be measured with arbitrary accuracy. As a matter of convention, the z-component is generally specified. The operator for the z-component of the angular momentum is denoted by L_z, and it also commutes with the Hamiltonian. Thus, the eigenvalues of L_z are also quantum numbers of the system. These eigenvalues are denoted by m_l, known as the magnetic quantum number. Whereas the total orbital angular momentum has magnitude $\hbar[l(l+1)]^{1/2}$, its projection along the z-axis has magnitude $m_l\,\hbar$.

When relativistic and spin effects are taken into account, atomic states characterized by the same n quantum number but different l quantum numbers are no longer degenerate. However, the collection of states with the same n and l quantum numbers remain degenerate. These states are further distinguished by the quantum number m_l which takes integer values from $-l$ to l, for a total of $2l+1$ degenerate states, each with quantum number n and orbital angular momentum quantum number l. These $2l+1$ states arise because an atom is degenerate with respect to its orientation in space (so-called space degeneracy). This degeneracy is eliminated in a magnetic field; the observed splitting of the spectral lines is known as the Zeeman effect.

In addition to orbital angular momentum, electrons have an intrinsic spin angular momentum, s, with magnitude $\frac{1}{2}\,\hbar$. A vector quantity, the intrinsic spin also has a direction. For electrons, the spin can either be "up" or "down." The direction of s gives rise to a fourth quantum number m_s, which takes values +1/2 and -1/2, corresponding to up and down respectively. Together, the four quantum numbers n, l, m_l, and m_s uniquely identify the stationary states of a hydrogenic atom.

2.1.2 Selection Rules for Hydrogenic Atoms

To see how quantum numbers are used in the labeling of spectral transitions, we will briefly examine the spectrum of the hydrogen atom. As explained above, the observed spectrum takes the form of a series of lines, each line corresponding to a particular wavelength of emitted radiation. The energy of the radiation is itself proportional to the frequency, ν', according to the formula $E_{radiation} = h\nu'$. Due to the conservation of energy, the energy of the emitted photon equals the change in energy of the atomic state:

$$\Delta E = E_{final\ state} - E_{initial\ state} = E_{radiation} = h\nu'$$

Thus, $\Delta E = h\nu'$. From this relationship between the frequency of the radiation and the spectral transitions, we can infer the energy levels of the hydrogen atom. Indeed, the observation of discrete spectral lines was some of the earliest evidence that the electronic states of atoms are quantized, with electrons only capable of occupying discrete states of energy, rather than a continuous range of energy values.[3]

These spectral transitions are organized using a Grotian diagram (also known as a term diagram or energy-level diagram), shown in Figure 1.

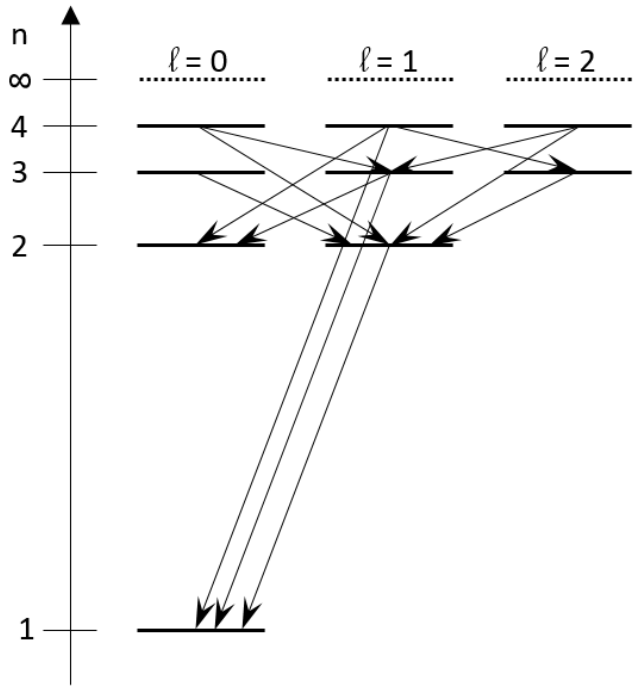

Figure 1. Grotian Diagram for Atomic Spectrum of Hydrogen

In Figure 1, the energy of the atomic states increases from the bottom to the top. The values of *n* run upwards in the diagram, forming rows. The values of *l* run across the diagram, forming columns. The lowest energy level is known as the ground state, which in the case of hydrogen has energy of -13.6 electron volts. As explained above, the energy of a given atomic state depends only on the quantum numbers *n* and *l* and is degenerate for the quantum numbers

[3] Upon ionization, the electron is no longer bound, and the atom exhibits a continuous spectrum. Selection rules only apply—and indeed only have sense—in the discrete portion of the atomic spectrum.

m_l and m_s. Thus, it is unnecessary to use m_l and m_s in labeling the diagram. The relative spacing of the energy levels shows that states with the same value of n, but different l values, have approximately the same energy, confirming the meaningfulness of the early neglect of relativistic and spin effects, which break this degeneracy.

Prima facie, it might be thought that given the various electronic energy levels, transitions could occur from any level to any other level. Yet after constructing an energy level diagram, we notice that spectral transitions occur between energy levels only if certain conditions are satisfied. These restrictions are known as *selection rules*, and they specify the possible states to which an electron in a given state can transition. For hydrogenic atoms transitioning under electric dipole radiation—the most intense radiative mechanism—the following selection rules are observed:

1. Δn is unrestricted. The principal quantum number can change by any integer value.
2. $\Delta l = \pm 1$. The orbital angular momentum quantum number can change only by 1.
3. $\Delta m_l = 0, \pm 1$. The magnetic quantum number can change by zero or one.

Returning to our simple example of the hydrogen spectra, we can see how these rules apply. For an electronic state with a given value of n, a spectral transition can occur to a state with any other value of n, provided that the other selection rules are satisfied. For a state with quantum number $l = 1$, spectral transitions can occur only to states with $l = 0$ or $l = 2$, provided again that the other rules are satisfied. Due to space degeneracy, the selection rule for the magnetic quantum number is not apparent on a Grotian diagram. This selection rule becomes apparent only when a hydrogen atom is placed within a magnetic field, which breaks a spatial symmetry and makes the states with different m_l distinguishable.

2.1.3 Stationary States and Selection Rules for Many-Electron Atoms[4]

For many electron atoms, the wavefunctions of the atom are treated to first approximation as a product of one-electron wavefunctions. In this so-called "orbital approximation," each electron moves in an approximate spherical potential created by the average interactions of all the other

[4] For further details, confer Herzberg (1944).

electrons. Since this potential is spherical, each electron can again be labeled with four quantum numbers n, l, m_l, and m_s. By labeling the electrons in this manner, we form electron configurations for the ground states of each atom. For instance, oxygen has an electron configuration in the ground state of $1s^2 2s^2 2p^4$, indicating that there are two electrons labeled with n=1 and l=0, two electrons with n=2 and l=0, and four electrons with n=2 and l=1.

In the many-electron case, the individual l_i's form a resultant L, which represents the total orbital angular momentum of the atom. Likewise, the individual s's for each electron are added to form a resultant S, representing the total spin angular momentum of the electrons orbiting the atom. These resultants L and S couple to form the total angular momentum, J, of an atom.

All three of these overall angular momentum quantum numbers L, S, and J are vector quantities and can take 2x+1 different directions along a preferred axis, which is arbitrarily chosen to be the z-axis. For instance, the z-component of the total angular momentum quantum number, M_J, is quantized in any one of 2J+1 possible directions, with magnitudes ranging over the values –J, –J + 1, . . ., J–1, J. Likewise, M_L ranges over –L, –L + 1, . . ., L–1, L and M_S over –S, –S + 1, . . ., S–1, S. Depending on the relative orientation of the total orbital angular momentum and the total spin angular momentum, a given electron configuration can exhibit a number of different energy states. These different energy states correspond to the spectroscopic terms organized empirically on an energy-level diagram. For instance, the term $^3P_2$ denotes an energy level characterized by a total orbital angular momentum of L=1 (denoted by P), and a spin multiplicity (denoted by the left superscript) of 3=2S+1, which implies a total spin angular momentum of 1. The orbital and spin angular momentum sum to a total angular momentum of J= 2, denoted by the right subscript.

For many-electron atoms, the selection rules are characterized in terms of the overall angular momentum quantum numbers L, S, and J. For spectral transitions arising from electric dipole radiation and occurring in atoms with atomic numbers less than 40, $\Delta S = 0$, $\Delta L = 0, \pm 1$, $\Delta J = 0, \pm 1$, and $\Delta M_J = 0, \pm 1$. The most intense transitions generally arise from the transition of a single electron. Thus, in addition to the selection rules for the resultant angular momenta, the selection rules seen in the hydrogenic case also hold. Namely, for the transitioning electron, $\Delta l = \pm 1$ and $\Delta m_l = 0, \pm 1$.

Transitions that violate these selection rules are called "forbidden" transitions, while transitions which satisfy all rules are "allowed" transitions. In reality, these selection rules are

only approximate restrictions that account for the most salient features of a system. Forbidden transitions that violate these rules are often weakly allowed, appearing as rather faint lines in the spectrum. A number of factors lead to violations of the selection rules, including higher order radiative transition mechanisms such as electric quadruple radiation and magnetic dipole radiation, the presence of electric or magnetic fields, coupling of the resultant angular momentum with the nuclear spin, and double-electron transitions. The selection rules listed above describe the restrictions on electric dipole radiation. This radiative mechanism is in general the most prominent, leading to lines 10^5 to 10^8 times more intense than those produced by higher order radiative mechanisms. For this reason, selection rules for electric dipole radiation are the ones most easily accounted for on energy-level diagrams.

For many electron atoms, the individual orbital and spin angular momenta (the l_i*'s* and *s's)* of the electrons provide only approximate descriptions of the atomic state. As the result of coupling of the orbital and spin angular momenta (known as spin-orbit coupling), the individual angular momenta lose their meaning as well-defined descriptions of separate momenta. This has the further consequence of making the resultant orbital and spin angular momenta, L and S, also approximate. For relatively light atoms, spin orbit coupling is small, and L and S provide useful descriptions of atomic states. In this approximation scheme, the individual l_i*'s* are coupled to form L, and the individual m_s*'s* coupled to form S. L and S are then coupled to form J. This is known as Russell-Saunder's (or LS) coupling, and it was used in formulating the selection rules above.

As the atomic number of the atom increases, the effects of spin-orbit coupling become more pronounced, and L and S lose their meaning. More frequent and intense transitions are seen which violate the $\Delta S = 0$ and $\Delta L = 0, \pm 1$ rules. For heavy atoms, a different coupling procedure known as j-j coupling is used. For mid-weight atoms, intermediate coupling schemes are most appropriate. Nevertheless, the total angular momentum J (and its z-component M_J) remains well-defined for all atoms. In the following, I focus on the selection rules for electric dipole radiation in atoms of low atomic number.

2.1.4 Matrix Elements of the Perturbation Operator

Early explanations of selection rules were given in terms of the correspondence principle, a frequent explanatory tool of the old quantum theory. With the advent of quantum mechanics, selection rules were soon derived as consequences of basic principles. A number of conceptually distinct derivations were developed, including wave-mechanical derivations employing the eigenfunctions of atomic Hamiltonians, matrix mechanical derivations using angular momentum operators, and group theoretical derivations utilizing irreducible representations of groups of commuting observables. A common goal unites these different approaches: the conditions must be found under which the matrix elements of the perturbation equal zero.

Spectral transitions are fundamentally time-dependent phenomena, and all derivations of selection rules begin with an application of time-dependent perturbation theory. Sparing the mathematical details, the result is that the probability of transition from an initial state ψ_i to a final state ψ_f is proportional to the squared modulus of a matrix element $\langle\psi_i|\boldsymbol{P}|\psi_f\rangle$, where $\boldsymbol{P}$ is the perturbation operator. The matrix element $\langle\psi_i|\boldsymbol{P}|\psi_f\rangle$ is equivalent to the integral over three dimensional space of the product of the complex conjugate of ψ_i with the product of $\boldsymbol{P}$ and ψ_f. Thus, we can write the following equality: $\langle\psi_i|\boldsymbol{P}|\psi_f\rangle = \int \psi_i^* \boldsymbol{P}\psi_f \, d\tau$. Both of these expressions are known as the transition moment integral (TMI), which always equals a real number.

Physically, the transition moment integral represents the interaction between the system and an external perturbation, such as impinging electromagnetic radiation. Under certain conditions (given by selection rules), this interaction causes the system to transition from an initial state (represented by ψ_i) to a final state. In emission spectroscopy, a transition results in the emission of radiation from the system.

The transition moment integral (TMI) is proportional to the intensity of emitted radiation. Thus, when the TMI equals zero, no transition occurs from the given initial to final states under the perturbation $\boldsymbol{P}$. Given this proportionality, the problem of deriving selection rules is equivalent to the problem of finding the conditions under which the transition moment integral equals zero. Since the most intense transitions arise from electric dipole radiation, I focus on the case where $\boldsymbol{P}$ is the operator for electric dipole radiation. This operator is given by $\boldsymbol{p} = -e_c\boldsymbol{r}$, where "$e_c$" stands for the elementary charge of an electron, and "$\boldsymbol{r}$" is the three-dimensional position vector; it has Cartesian components $\hat{x}$, $\hat{y}$, and $\hat{z}$. Insofar as we are interested in

accounting for the most salient spectral transitions, we are mainly interested in those arising from an electric dipole mechanism.

2.1.5 Selection Rules as an Example of Universality

In the remainder of this chapter, I present derivations of selection rules from both the commutator and group theoretic viewpoints. As we examine these derivations, it is important to keep in mind what we are trying to explain. We are not particularly interested in explaining why a particular atomic system exhibits the selection rules that it does in a given experimental setup. Rather, we are primarily interested in explaining why a large class of atomic systems exhibits the same selection rules. Batterman (2002) has developed an account of explanation tailored to these sorts of shared patterns of behavior, which are called universal behaviors. In this section, I argue that selection rules fall under the scope of Batterman's account. I return to this point in Chapter 3 where I consider Batterman and Rice's (unpublished) account of minimal model explanations, which explain universal behaviors. Batterman and Rice's account bears directly on the case study that I examine here.

Selection rules are an example of a particular kind of physical phenomenon, known as a universality behavior. Physicists describe behavior as universal when many, structurally distinct, physical systems exhibit analogous behavior under analogous circumstances. As a paradigmatic example, we can consider phase transitions exhibited by fluids and magnets at their critical points. Even though these systems have vastly different micro-physical constitutions, many of them undergo phase transitions according to the same critical exponent (Batterman 2002, chapter 4). Similarly, atomic systems with low atomic numbers (less than 40 protons in the nucleus) exhibit the same set of selection rules. These atomic systems are micro-structurally distinct, possessing different numbers of protons, neutrons, and electrons along with different electron configurations. Nevertheless, upon excitation by electromagnetic radiation, they emit spectral lines governed by the same pattern of selection rules. These spectral lines are themselves a macroscopic phenomenon. Some spectral lines—such as those of hydrogen's Balmer series—even occur in the visible spectrum. Although we only notice selection rules when we reorganize the frequencies of these spectral lines into an energy level diagram, it is fair to say that selection

rules govern a macroscopic phenomenon. They determine which spectral lines are allowable by constraining transitions between atomic energy levels.

To explain a case of universality, we embark upon a particular kind of explanatory task. In his account of scientific explanation, Batterman (2002) characterizes explanations of universality as a species of type (ii) why-questions. A type (ii) why-question "asks why, in general, patterns of a given type can be expected to obtain" (Batterman 2002, p. 23). Explanations of universality are of this form. We are interested in explaining why a class of disparate systems displays a shared behavior. Batterman distinguishes type (ii) why-questions from a distinct explanatory task, termed type (i) why-questions. Type (i) why-questions probe why a given physical system exhibited the behavior it did under a given circumstance (2002, p. 23). For instance, if I am working on a physical chemistry lab experiment, I might be interested in knowing why my particular system of hydrogen atoms exhibited the spectrum it did on that day. Answering this question motivates a certain explanatory task: along with more general theoretical considerations, I would focus on details of my apparatus and the state of my system. Yet, in explanations of atomic behavior in general, we are generally interested in type (ii) why-questions. Speaking about theoretical atomic spectroscopy, Judd remarks that, "The purpose of the analysis is not just to account for the properties of a particular atom as closely as possible but also to gain insight into its structure and to discern features that are shared by other atomic systems" (1963, p. 1). Some of these type (ii) why-questions—such as those dealing with selection rules—are specifically questions about the universality of a behavior displayed by atomic systems.

2.2 COMMUTATOR APPROACHES TO SELECTION RULES[5]

A framework for computing matrix elements—and hence for deriving selection rules—is provided by the use of fundamental commutation relationships. As in classical mechanics, the angular momentum is the vector cross product of the position and momentum vectors:

[5] Confer Griffiths (2005), Chapters 4 and 9, for additional details.

$$\boldsymbol{J} = \boldsymbol{r} \times \boldsymbol{p}.$$

In terms of the Cartesian components, this equation means the following:

$$J_x = yp_z - zp_y; \qquad J_y = zp_x - xp_z; \qquad J_z = xp_y - yp_x$$

These equations define the components of the angular momentum as functions of the position and momentum. The position and momentum operators are related to each other through the canonical commutation relations:

$$[r_i, p_j] = -[p_i, r_j] = i\hbar\delta_{ij} \qquad [r_i, r_j] = [p_i, p_j] = 0\,.$$

These equations show that the only position and momentum operators that fail to commute are x with p_x, y with p_y, and z with p_z. The fundamental commutation relations for angular momentum follow from these relationships:

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k$$

where ε_{ijk} is the Levi-Civita symbol, which equals 1 for an even permutation of i, j, k (which represent the x, y, and z components), –1 for an odd permutation, and zero if any index is repeated.

Because different components of angular momentum do not commute with each other, it is impossible to simultaneously measure the angular momentum in any two directions. However, the square of the total angular momentum, J^2 commutes with each component.

$$J^2 \equiv J_x^2 + J_y^2 + J_z^2 \qquad [J^2, J_i] = 0$$

As explained earlier, since J^2 and J_z commute, it is possible to measure the square of the total angular momentum and the z-component of the angular momentum simultaneously. In those physical situations where the orbital angular momentum is well-defined, identical commutation relations hold for L^2, L_x, L_y, and L_z.

2.2.1 Commutation Derivation of the Selection Rule for Δm_l

The commutator approach to angular momentum shows that selection rules are a consequence of the following theorem and its extensions. Within the context of selection rules, this theorem states that if an operator $\hat{T}$ commutes with the perturbation operator, then the matrix element $\langle\psi_i|\boldsymbol{P}|\psi_f\rangle$ vanishes unless the initial and final stationary states share the same eigenvalue of $\hat{T}$.

Theorem: Let $\hat{T}$ be a Hermitian operator such that $\hat{T}$ and $\hat{P}$ commute, i.e. $[\hat{T}, \hat{P}] = 0$. Let ψ_i and ψ_f be eigenfunctions of $\hat{T}$ associated with eigenvalues t_i and t_f respectively. Then $\langle\psi_i|\boldsymbol{P}|\psi_f\rangle = \int_0^\infty \psi_i^* \boldsymbol{P} \psi_f \, d\tau = 0$ unless $t_i = t_f$.

Proof:

Since $[\hat{T}, \hat{P}] = 0$, it follows that $\langle\psi_i|[\hat{T}, \hat{P}]|\psi_f\rangle = 0$.

Expanding the commutator, $\langle\psi_i|[\hat{T}, \hat{P}]|\psi_f\rangle = \langle\psi_i|\hat{T}\hat{P} - \hat{P}\hat{T}|\psi_f\rangle = 0$

Since $\hat{T}$ is Hermitian and ψ_i and ψ_f are eigenfunctions of $\hat{T}$,

$$\langle\psi_i|\hat{T}\hat{P} - \hat{P}\hat{T}|\psi_f\rangle = \langle\psi_i|t_i\hat{P} - t_f\hat{P}|\psi_f\rangle = 0$$

$$= (t_i - t_f)\langle\psi_i|\hat{P}|\psi_f\rangle = 0$$

Thus, either $t_i = t_f$ or $\langle\psi_i|\hat{P}|\psi_f\rangle = 0$. ■

This theorem supplies a general selection rule for eigenvalues of operators that commute with the perturbation operator. From this theorem, the selection rule for m_l follows directly.

The above commutation relations imply that L_z and z commute, i.e. $[\hat{L}_z, z] = 0$. As explained earlier, the eigenvalues of $\hat{L}_z$—the operator for the z-component of the angular momentum—are the magnetic quantum numbers, m_l. Furthermore, the operator for radiation polarized in the z-direction is itself proportional to z. Thus, from this commutation relation and the theorem above, it follows that $\langle\psi_i|\hat{z}|\psi_f\rangle = 0$ unless $m_i = m_f$, i.e., unless $\Delta m_l = 0$.

Although, $\hat{L}_z$ does not commute with $\hat{x}$ or $\hat{y}$, the selection rule for radiation polarized in the x or y direction can be derived in an analogous way. The commutator $[\hat{L}_z, x] = i\hbar y$ is placed between the wavefunctions of the initial and final states:

$$\langle\psi_i|[\hat{L}_z, x]|\psi_f\rangle = \langle\psi_i|\hat{L}_z x - \hat{L}_z x|\psi_f\rangle = (m_i - m_f)\hbar\langle\psi_i|x|\psi_f\rangle = i\hbar\langle\psi_i|y|\psi_f\rangle$$

Thus, $(m_i - m_f)\langle\psi_i|x|\psi_f\rangle = i\langle\psi_i|y|\psi_f\rangle$ (*)

Likewise, using the commutator, $[\hat{L}_z,\ y] = -i\hbar x$ yields

$$\langle\psi_i|[\hat{L}_z, y]|\psi_f\rangle = \langle\psi_i|\hat{L}_z y - \hat{L}_z y|\psi_f\rangle = (m_i - m_f)\hbar\langle\psi_i|y|\psi_f\rangle = -i\hbar\langle\psi_i|x|\psi_f\rangle$$

So, $(m_i - m_f)\langle\psi_i|y|\psi_f\rangle = -i\langle\psi_i|x|\psi_f\rangle$ (**)

Multiplying (*) by $(m_i - m_f)$, yields $(m_i - m_f)^2\langle\psi_i|x|\psi_f\rangle = i(m_i - m_f)\langle\psi_i|y|\psi_f\rangle$

Multiplying (**) by i yields $i(m_i - m_f)\langle\psi_i|y|\psi_f\rangle = \langle\psi_i|x|\psi_f\rangle$.

Thus, $(m_i - m_f)^2\langle\psi_i|x|\psi_f\rangle = \langle\psi_i|x|\psi_f\rangle$.

It follows that either $(m_i - m_f)^2 = 1$ or the matrix element $\langle\psi_i|x|\psi_f\rangle$ vanishes. Thus, the selection rule for radiation polarized in the x-direction is $\Delta m_l = \pm 1$. An analogous derivation shows that the same selection rule holds for radiation polarized in the y-direction.

2.2.2 Commutator Derivation of the Selection Rule for Δl

A commutator derivation of the rule $\Delta l = \pm 1$ was found early on during the development of matrix mechanics, appearing first in a paper of Born, Heisenberg, and Jordan from 1926. Griffiths' (2005) presentation utilizes the following commutator relationship:

$$[L^2, [L^2, r]] = 2\hbar^2(rL^2 + L^2 r)$$

This commutation relation is derivable from the above commutation relations, but the derivation is not trivial. This commutation relation implies that

$$\langle l'm'|[L^2, [L^2, r]]|lm\rangle = \langle l'm'|2\hbar^2(rL^2 + L^2 r)|lm\rangle$$

The right hand side equals $2\hbar^4\big(l'(l'+1) + l(l+1)\big)\langle l'm'|r|lm\rangle$.

Evaluating the left hand side yields $\langle l'm'|L^2[L^2, r] - [L^2, r]L^2|lm\rangle$

$$= \hbar^2\big(l'(l'+1) - l(l+1)\big)\langle l'm'|[L^2, r]|lm\rangle$$

$$= \hbar^2\big(l'(l'+1) - l(l+1)\big)\langle l'm'|L^2 r - rL^2|lm\rangle$$

$$= \hbar^4\big(l'(l'+1) - l(l+1)\big)^2\langle l'm'|r|lm\rangle$$

Subtracting the left hand side from the right hand side yields

$$0 = \langle l'm'|r|lm\rangle \left\{2\hbar^4\big(l'(l'+1)+l(l+1)\big) - \hbar^4\big(l'(l'+1)-l(l+1)\big)^2\right\}.$$

Thus, either $\langle l'm'|r|lm\rangle = 0$ or $2\hbar^4\big(l'(l'+1)+l(l+1)\big) = \hbar^4\big(l'(l'+1)-l(l+1)\big)^2$.

With regards to the left hand side of the second equation above, note that

$$2\big(l'(l'+1)+l(l+1)\big) = (l'+l+1)^2 + (l'-l)^2 - 1.$$

Regarding the right hand side, note that $l'(l'+1) - l(l+1) = (l'+l+1)(l'-l)$.

These equations imply that

$$(l'+l+1)^2 + (l'-l)^2 - 1 = (l'+l+1)^2(l'-l)^2$$

which in turn implies that

$$(l'+l+1)^2(l'-l)^2 - (l'+l+1)^2 - (l'-l)^2 + 1 = 0$$
$$= [(l'+l+1)^2 - 1][(l'-l)^2 - 1] = 0.$$

Thus, either $(l'+l+1)^2 - 1 = 0$ or $(l'-l)^2 - 1 = 0$. $(l'+l+1)^2 - 1$ can equal zero only if $l' = -l$. Since the angular momentum quantum numbers are always greater than or equal to zero, $l' = -l$ only if $l' = l = 0$. This possibility is in fact disallowed by a selection rule on parity (Laporte's Rule), which states that in a one-electron transition, the orbital angular momentum must change from an odd value to an even value. This rule can be derived via a parity operator.

This leaves the second equation: the expression $(l'-l)^2 - 1$ can equal zero only if $l' = l \pm 1$. Thus, the selection rule for Δl is obtained.

2.2.3 Generalized Commutator Approach

Although the above derivation for the Δl rule is valid, it fails to motivate the initial commutator relation

$$\big[L^2, [L^2, r]\big] = 2\hbar^2(rL^2 + L^2 r).$$

Condon and Shortley employ a similar derivation for the selection rule on J, the total angular momentum. At the beginning of their derivation for the rule on Δj, Condon and Shortley reference the 2nd edition (1935) of Dirac's *Principles of Quantum Mechanics* as the source of a

general method for deriving selection rules. Dirac—already in the first edition (1930) of his text—provides a general characterization of the selection rule problem from the standpoint of commutator algebra.

Dirac explains that in order to derive a selection rule for stationary states—characterized by a complete set of commuting observables A—under a perturbation P, one must find an algebraic equation that is linear in the perturbation operator and is a function only of P and the operators of the set A. Dirac writes the general form of this equation as

$$\sum_r f_r P g_r = 0 \tag{1}$$

where the functions f_r and g_r depend only on the operators of the set A.

In a particular representation (e.g. Schrödinger, Heisenberg, etc.), f_r and g_r are explicit functions of the eigenvalues of the operators A, and P can be written in terms of its matrix elements. By focusing on an operator T of the set A, with eigenvalues *t* and *t'* for the initial and final states, we can write (1) above as:

$$\sum_r f_r(t)\langle t|P|t'\rangle g_r(t') = 0$$

$$\langle t|P|t'\rangle \sum_r f_r(t) g_r(t') = 0$$

Unless the expression $\sum_r f_r(t) g_r(t')$ vanishes, the matrix element $\langle t|P|t'\rangle$ necessarily equals zero. Hence, if an algebraic equation of the form (1) can be found, a selection rule for the operator T under perturbation P follows.

This general framework motivates the derivation of the commutator equation $\left[L^2,[L^2,r]\right] = 2\hbar^2(rL^2 + L^2 r)$. Earlier, in deriving the selection rule for Δm_l, we formed the commutator $[L_z,\ z] = 0$, which yielded an equation of the desired form (1). Since $l(l+1)$ is the eigenvalue of the total angular momentum operator L^2, we proceed by forming $[L^2,\ z]$, which upon evaluation equals $2i\hbar\left(xL_y - yL_x - i\hbar z\right)$.

However, this expression is not of the form (1), for it depends on perturbation operators x, y, and z, rather than a single vector component. Furthermore, it contains operators L_y, and L_x, but the goal is to have f_r and g_r be functions of L^2. If f_r and g_r are functions of L^2, then (1) provides an algebraic restriction on the eigenvalues of L^2.

Similar problems arise when we evaluate $[L^2,\ x]$ and $[L^2,\ y]$, which (under cyclic permutation of x, y, z) equal the following:

$$[L^2, \mathrm{x}] = 2i\hbar\left(yL_z - zL_y - i\hbar x\right)$$

and

$$[L^2, \mathrm{y}] = 2i\hbar(zL_x - xL_z - i\hbar y)$$

Noting that $L^2 \equiv\ L_x^2 + L_y^2 + L_z^2$, we are led to try to find a commutator that leads to this sum. Forming the more complicated commutator $\left[L^2, [L^2, z]\right]$ fulfills this desideratum. Ultimately, it is found that $\left[L^2, [L^2, z]\right] = 2\hbar^2(zL^2 + L^2 z)$, which can be generalized to a commutator expression involving *r*. Expanding the commutator on the left hand side results in $L^4 r - 2L^2 r L^2 + r L^4 - 2\hbar^2(rL^2 + L^2 r) = 0$, which is of the form (1).

Thus, although neither Dirac nor Condon and Shortley explicitly motivate their use of this commutation equation, its use is not entirely unmotivated. Dirac proceeds to derive the selection rule for Δl (although he uses a modified definition of the total angular momentum operator). He then generalizes this derivation to find the selection rule for the total angular momentum.

Condon and Shortley eschew Dirac's reliance on the Cartesian components of the perturbation operator and derive the rule for Δj via the vector expression of the perturbation. Nevertheless, the differences between the derivations of Dirac, Condon and Shortley, and the one employed above are little more than notational variants. At the heart is the commutator relation for $\left[L^2, [L^2, r]\right]$, from which the selection rule follows.

2.3 GROUP THEORETIC APPROACHES TO SELECTION RULES[6]

Due to the complexities of group theory and group representation theory, I focus on details required for understanding the group theoretic approach to selection rules. In simple terms, groups are symmetries of geometrical figures, such as triangles or squares. An equilateral triangle possesses six symmetries: three rotations and three reflections. If an equilateral triangle is rotated by 120°, 240°, or 360°, the resultant figure looks the same: the triangle looks as though nothing was done to it. This invariance property characterizes these three rotations as symmetries of equilateral triangles. Rotating the triangle by 60° would result in a noticeably different figure, with the vertices of the triangle shifted. Likewise, each perpendicular bisector running from a vertex to the opposite side provides a mirror plane: reflecting all the points of the triangle over this line results in an identical triangle. Equilateral triangles have three perpendicular bisectors with this property, providing three reflection symmetries. Together, these six symmetries—the three rotations and three reflections—form the symmetry group of an equilateral triangle.

In quantum mechanical situations, we are interested in the symmetry group of the Hamiltonian operator. The Hamiltonian is a mathematical construct encoding information about a physical system's energy states. Since in spectroscopy we are concerned with transitions between energy states, it is not surprising that the Hamiltonian plays an important role. Just as a group of geometrical symmetry operations leave an equilateral triangle invariant, there is a group of quantum mechanical operators that leave the Hamiltonian invariant. This set of operators is known as the group of the Hamiltonian, or alternatively, as the group of the Schrödinger equation. What operators belong to this group depends on the particular physical system under consideration. In this case study, we are largely concerned with an idealization of atomic systems known as the central field approximation. In this idealization, we average electronic repulsions as a "central field," yielding a potential energy function that depends only on the radial distance of an electron from the nucleus. This potential energy has spherical symmetry: an electron at a given radial distance sees the same potential no matter its angular coordinates. In other words, the potential energy function—and furthermore the Hamiltonian—is invariant under three-

[6] For further details, confer Tinkham (1964) and Petrashen and Trifonov (1969).

dimensional rotations. These rotations form the group of the Hamiltonian for this idealized atomic system.

Mathematically, it is cumbersome to work with geometrical symmetry operations. Although these are advantageous for visualization, they present computational difficulties. Instead, it is convenient to represent these geometrical symmetry operations with matrices that preserve the multiplicative relations of the original geometrical operations. For instance, if the outcome of a rotation A followed by a rotation B is the same as an overall rotation C, then the matrices representing A and B should multiply in the same way, yielding a matrix that represents rotation C. Symbolically, if $B \cdot A = C$, then $R(B) \cdot R(A) = R(C)$, where *R(B)* is a matrix representing the rotation B. A matrix representation of a group is a set of matrices that preserve the multiplicative properties of the original group. Matrix representations are powerful because they enable us to analyze a geometrical situation in terms of linear algebra. This makes a number of tractable solution strategies immediately applicable. Thoughtfully unpacking this ascription of intellectual power would require a story like the one I develop in Chapter 4 for derivations of selection rules.

In fact, for any given group, there are an infinite number of nonequivalent matrix representations. Fortunately, each representation decomposes into a manageable number of basic building blocks, known as irreducible representations. These irreducible representations encode the underlying symmetry types of a physical situation. They specify how functions transform under operations from a given symmetry group. Underlying the application of group theory to quantum mechanics is the following fact: stationary states of the atom (eigenfunctions of the Hamiltonian) transform as irreducible representations of the group of the Hamiltonian. This means that each stationary state wavefunction can be labeled by an irreducible representation that specifies its symmetry type. Arbitrary functions do not possess this property. In general, functions will transform as non-trivial reducible representations, which can be decomposed into sums of irreducible representations.

2.3.1 Selection Rules from a Group-Theoretical Perspective

Group theory partially re-expresses the matrix element $\langle \psi_i | \boldsymbol{P} | \psi_f \rangle$ in terms of the symmetry of the unperturbed atomic system. Regarding the restricted case of selection rules, group theory

completely re-expresses the necessary—but not sufficient—conditions for non-zero matrix elements.

As before, the problem is to determine when the matrix element $\langle\psi_i|\boldsymbol{P}|\psi_f\rangle$ necessarily vanishes. Denote the symmetry group of the unperturbed quantum mechanical system by G. Group representation theory enables a formal expression of the symmetry of a quantum mechanical operator and the symmetry of a wavefunction. As a simple example, consider a group theoretical recasting of the earlier theorem in Section 2.2.1 stating that if the perturbation operator $\hat{P}$ commutes with the operator $\hat{T}$—where ψ_i and ψ_f are eigenfunctions of $\hat{T}$—then the matrix element $\langle\psi_i|\hat{P}\,|\psi_f\rangle$ vanishes unless ψ_i and ψ_f are associated with the same eigenvalue of $\hat{T}$, i.e. $t_i = t_f$.

Consider first the special case where the perturbation operator $\hat{P}$ is invariant under *all* operations of the group G. Then $\hat{P}$ transforms as the trivial representation of G, denoted by Γ^0. In the context of matrix representations, the trivial representation Γ^0 maps every group element of G to a one-dimensional identity matrix. Thus, the traces of the matrices of this representation—known as the characters of the representation—all equal 1. Let the wavefunctions of the initial and final states transform as the irreducible representations denoted by Γ^i and Γ^f. Then the function $\hat{P}\psi_f$ transforms as the direct product representation $\Gamma^0 \otimes \Gamma^f$, which simply equals Γ^f because the characters of Γ^0 are 1 for all group elements of G. Hence, it follows from the orthogonality of basis functions that unless ψ_i and ψ_f belong to the same irreducible representation (more specifically, the same row of the same irreducible representation), the matrix element $\langle\psi_i|\hat{P}\,|\psi_f\rangle$ vanishes.

In the general case where the perturbation operator transforms by a more complex symmetry of G, group representation theory leads to the following criterion for non-vanishing matrix elements. Let the initial state ψ_i transform as the irreducible representation given by Γ^i, the final state ψ_f transform as Γ^f, and the perturbation operator $\hat{P}$ transform as Γ^p. Here, Γ^i and Γ^f are irreducible representations of the symmetry group G of the Hamiltonian of a given physical system. In the general case, Γ^p is a reducible representation of G.

Theorem: If the triple direct product $\Gamma = (\Gamma^i)^* \otimes \Gamma^p \otimes \Gamma^f$ does not contain the trivial irreducible representation, Γ^0, of G, then the matrix element $\langle\psi_i|\hat{P}\,|\psi_f\rangle$ necessarily vanishes.

Proof:

To say that ψ_i transforms as the irreducible representation given by Γ^i means that

$$T_g\psi_i = \sum_j \Gamma^i_{jk}(g)\psi_j.$$

Likewise, ψ_f transforms as Γ^f:

$$T_g\psi_f = \sum_m \Gamma^f_{ml}(g)\psi_m.$$

To say that the perturbation operator $\hat{P}$ transforms as Γ_P means that

$$T_g P_a {T_g}^{-1} = \sum_{b=1}^{k} \Gamma^p_{ba}(g) P_b$$

T_g is a unitary operator, so it preserves inner products. Thus, the matrix element equals:

$$\langle\psi_i|\hat{P}|\psi_f\rangle = \langle T_g\psi_i|T_g\hat{P}|\psi_f\rangle = \langle T_g\psi_i|T_g\hat{P}{T_g}^{-1}T_g|\psi_f\rangle$$

Substituting the corresponding summation expressions for $T_g\psi_i, T_g\psi_f, and\ T_g\hat{P}{T_g}^{-1}$ yields:

$$\langle\psi_i|\hat{P}\,|\psi_f\rangle = \left\langle \sum_j \Gamma^i_{jk}(g)\psi_j. \middle| \sum_{b=1}^{k} \Gamma^p_{ba}(g) P_b \middle| \sum_m \Gamma^f_{ml}(g)\psi_m \right\rangle$$

$$= \sum_{j,b,m} \Gamma^i_{jk}(g)\Gamma^p_{ba}(g)\,\Gamma^f_{ml}(g)\langle\psi_j|P_b|\psi_m\rangle$$

For brevity, I denote the product of the representation elements $\Gamma^i_{jk}(g)\Gamma^p_{ba}(g)\Gamma^f_{ml}(g)$ by $\Gamma_{jbm,\ iaf}$. These components form a matrix Γ equal to the direct product of the representations $(\Gamma^i)^*\otimes\Gamma^p\otimes\Gamma^f$. The equation above can be summed over the N group elements of the group G, yielding:[7]

$$\langle\psi_i|\hat{P}\,|\psi_f\rangle = \frac{1}{N}\sum_g \sum_{j,b,m} \Gamma_{jbm,\ iaf}\,\langle\psi_j|P_b|\psi_m\rangle \qquad (2)$$

The matrix $\Gamma = (\Gamma^i)^*\otimes\Gamma^p\otimes\Gamma^f$ can be decomposed into a sum of irreducible representations of the group G by a similarity transformation under the unitary matrix V:

[7] In cases where G is a linear Lie group, we replace these sums with integrals over group space.

$$V\,\Gamma(g)\,V^{-1} = \sum_{\lambda}^{\oplus} \Gamma^{\lambda}(g)$$

where the summation is over all irreducible representations of G. Hence,

$$\sum_{g} \Gamma(g) = V^{-1} \sum_{g} \sum_{\lambda}^{\oplus} \Gamma^{\lambda}(g)\,V$$

It is an important fact that the sum over the group G of the components of all irreducible representations equal zero unless the irreducible representation is the trivial one, Γ^0:

$$\sum_{g} \Gamma^{\lambda}(g) = 0 \text{ unless } \Gamma^{\lambda} = \Gamma^{0}$$

It follows that if the trivial representation is not one of the summands in the decomposition of $\Gamma = (\Gamma^i)^* \otimes \Gamma^p \otimes \Gamma^f$, then $\sum_g \sum_{\lambda}^{\oplus} \Gamma^{\lambda}(g)$ equals zero. It follows from (2) above that unless this decomposition contains Γ^0, the matrix element $\langle \psi_i | \hat{P} | \psi_f \rangle$ necessarily equals zero. ■

The application of this theorem to particular problems depends on knowledge of the irreducible representations for different symmetry groups. Atoms have approximate spherical symmetry, formally expressed by the three dimensional orthogonal group O(3). O(3) is the group of all rotations and reflections in three dimensional space $\mathbb{R}^3$. By expressing the transformation properties of atomic wave functions and the electric dipole operator in terms of irreducible representations of the orthogonal group, the selection rules for angular momentum quantum numbers are easily derived.

2.3.2 Group Theoretic Derivation of the Selection Rule for Δl

Once again, recall that electric dipole radiation is represented by the quantum mechanical operator $e\boldsymbol{r}$, where e is in this case the elementary charge of an electron. The radial vector $\boldsymbol{r}$

belongs to the irreducible representation Γ^1 of the orthogonal group. As before, denote the angular momentum of the initial state by l_i and the angular momentum of the final state by l_f. As explained in Section 2.3.1, the wave function of each electron must belong to an irreducible representation of the rotation group. Let Γ^{l_i} and Γ^{l_f} denote the irreducible representations that the initial and final wavefunctions, respectively, transform as.

It follows from the theorem above that to determine when the matrix element $\langle\psi_i|\boldsymbol{r}|\psi_f\rangle$ must necessarily vanish, it is sufficient to analyze the conditions under which the direct product representation $\Gamma = (\Gamma^{l_i})^* \otimes \Gamma^p \otimes \Gamma^{l_f} = (\Gamma^{l_i})^* \otimes \Gamma^1 \otimes \Gamma^{l_f}$ fails to contain the trivial representation Γ^0 in its decomposition.

The Clebsch-Gordan formula entails that $\Gamma^1 \otimes \Gamma^{l_i} = \Gamma^{l_i+1} + \Gamma^{l_i} + \Gamma^{l_i-1}$. For $\Gamma^{l_f} \otimes \Gamma^1 \otimes \Gamma^{l_i}$ to contain Γ^0, Γ^{l_f} must be one of the irreducible representations obtained in the direct product of $\Gamma^1 \otimes \Gamma^{l_i}$. Thus, Γ^{l_f} must equal Γ^{l_i+1}, or Γ^{l_i}, or Γ^{l_i-1}. Thus, l_f must equal $l_i + 1$, l_i, or $l_i - 1$. This implies that $\Delta l = \pm 1$ or 0.

As in the commutator derivation of the Δl selection rule, the possibility that $\Delta l = 0$ is blocked by considerations of parity. Laporte's rule states the additional requirement that the parity of an atom must change during a spectral transition induced by electric dipole radiation. While the commutator approach uses the parity operator to derive this rule, the group theoretic derivation relies on a function's evenness or oddness. Even functions are invariant under inversion, while odd functions change sign under inversion. The radial vector $\boldsymbol{r}$ is an odd function, and thus transforms as an odd representation of the orthogonal group. A one-electron wavefunction is even if its orbital angular momentum l is even; it is odd if l is odd. Group theoretically, it is clear that $(\Gamma^{l_i})^* \otimes \Gamma^1 \otimes \Gamma^{l_f}$ can transform as Γ^0—an even representation—only if the parities of the initial and final wavefunctions are different, e.g. if the initial state transforms as an even representation, then the final state must transform as an odd representation.

2.4 SUMMARY

Comparing the commutator derivations of Sections 2.2.1 and 2.2.2 with the group theoretic derivations of Section 2.3.2, it is clear that these approaches operate on distinct planes of thought. Although both approaches are motivated by the strategy outlined in Section 2.1.4—namely, determine necessary conditions for matrix elements of the perturbation operator to be zero—they exhibit different realizations of this strategy. As exemplified by Dirac's procedure (Section 2.2.3), the commutator approach remains closely tied to the algebra of quantum mechanical operators. It implements the matrix element strategy of Section 2.1.4 in a more direct manner than the group theoretic approach. One advance enabled by group theory is the reinterpretation of matrix elements in terms of their symmetries. The symmetry constraint derived in Section 2.3.1 provides a different strategy for determining when matrix elements are necessarily zero. Expressed in terms of the direct products of irreducible representations, this symmetry constraint appears conceptually distant from the original strategy, which is expressed directly in terms of operators. Nonetheless, it provides a more tractable path for deriving selection rules.

In Chapter 4, I undertake a thorough analysis of expressive differences between the two approaches. I argue that group theory's re-expression—and subsequent reorganization—of constraints governing spectral transitions provides a deeper understanding of selection rules. Before developing my account of scientific understanding, I consider whether these derivations provide explanations of selection rules. From the arguments presented above, it is clear that these approaches justify selection rules. Nevertheless, these derivations may be explanatorily deficient, disqualifying them from being genuine explanations. It is to this issue that I now turn.

3.0 EXPLAINING SELECTION RULES IN ATOMIC SPECTROSCOPY

In this chapter, I provide an answer to the question, "In virtue of what criteria are the derivations of selection rules explanatory?" On some philosophical accounts of scientific explanation, at least some of the derivations in Chapter 2 justify selection rules but nevertheless fail to explain them. I consider two such accounts in Sections 3.2 and 3.4, respectively. I argue that these accounts do not (at least not always) accurately describe scientific practice. I present an alternative account of mathematical scientific explanations in Section 3.1. My use of the word "explanation" is deflationary: it is considerably less rich than usual philosophical notions of "genuine explanation." On my account, an argument is explanatory if it recovers the phenomenon of interest according to principled mathematical and physical constraints. My account does not distinguish derivations from explanations. Instead, I treat the usual richness of the word "explanation" under the heading "understanding." I believe my way of speaking is more felicitous for an analysis of how mathematics advances physical insight.

My account of explanation is at odds with an alternative account criticized by Robert Batterman and Collin Rice (unpublished). Batterman and Rice refer to this family of positions as common features accounts. According to common features accounts, derivations may justify that a phenomenon occurs, but unless a derivation brings out the relevant features, it cannot be explanatory. In Section 3.2, I introduce philosophical positions that defend this viewpoint, which I refer to as relevance accounts of explanation. In Section 3.3, I raise two problems for relevance accounts of explanation, both epistemic in nature. Section 3.3.1 argues that relevance accounts are circular. I contend that we often determine what physical and mathematical details are relevant only by constructing a successful explanation—in my deflated sense of explanation. A successful mathematical explanation justifies these details as relevant and neglected details as irrelevant. I argue that if my weaker notion of explanation grounds ascriptions of relevance, then

relevance should not be used as a criterion for explanation. Section 3.3.2 develops a related problem stemming from the existence of multiple explanations of a given phenomenon.

Section 3.4 considers the minimal model account of explanation proposed by Batterman and Rice. Batterman and Rice raise similar problems for relevance accounts, and their account of relevance avoids the problems I develop. Nevertheless, in Section 3.5 I argue that if taken as a requirement for explanation, their methodological prescriptions are too strong. I argue for a weaker interpretation where these methodological prescriptions are suggestions—rather than requirements—for good explanatory practices. In doing so, I defend my nonstandard account of explanation from criticisms supplied by these more restrictive accounts.

3.1 A DEFLATIONARY ACCOUNT OF EXPLANATION

In this section, I present my account of scientific explanations that involve mathematics.[8] As noted previously (confer Section 1.1), I use "explanation" in a nonstandard sense. For me, "explanation" is a minimal requirement for an argument to provide understanding. I bundle much of the richness of philosophers' usual conception of explanation under the term "scientific understanding," which I examine in the next chapter. My nonstandard account of explanation proposes two minimal criteria for an argument to be explanatory: (1) recovering the phenomenon to be explained and (2) doing so according to accepted mathematical and physical constraints. I argue that these two criteria are sufficient for a deflated notion of scientific explanation.

First, an explanatory argument must recover the phenomenon to be explained. To "recover a phenomenon," an argument must show how this phenomenon is the endpoint of a chain of reasoning. At the end of the argument, we arrive at the phenomenon of interest. For instance, derivations of selection rules terminate with a selection rule; it is in this sense that they "recover a selection rule." If an argument fails to recover the phenomenon of interest (either approximately or exactly) then it is a failed argument and has no chance of being explanatory. We begin to think that an argument might be explanatory when it takes us to the point we wish to

[8] I do not intend my account of scientific explanation to be wholly general.

reach, i.e. it recovers the explanandum. Hence, recovery of the phenomenon of interest is a minimal requirement for explanation.

Second, an explanatory argument must proceed according to principled reasons. We cannot reason in any fashion we deem suitable. Explanatory arguments must recover the explanandum as a consequence of accepted mathematical and physical constraints. Examples of physical constraints include boundary conditions, restrictions on the range of possible measurement outcomes, and accepted physical laws. Mathematical constraints include theorems, methods of computation, algebraic identities, and solution procedures. In using the phrase "accepted constraints," it may seem that I should provide criteria for acceptance. I am skeptical that general criteria can be provided that accurately describe the varied conditions under which scientists and mathematicians agree on constraints.

My account has similarities with Hempel and Oppenheim's deductive-nomological model of explanation (1965). As in Hempel's account, I take sufficiently constrained derivations that recover the explanandum to be explanatory. I identify these kinds of derivations with explanations. One problem with such accounts is that they seem to sanction too many arguments as explanatory. This worry motivates separate restrictions supported by relevance accounts and by Batterman and Rice. Faced with multiple purported explanations, these accounts attempt to pick out the genuine explanation, which is generally taken to be singular. This way of talking leads to difficulties when analyzing multiple explanations of the type considered below in Section 3.3.2. Faced with multiple compatible explanations of a given phenomenon, we need a way to distinguish them based on the insights they provide. At the same time, it is natural to require that only explanatory arguments can provide insights. One strategy—the strategy presumably supported by relevance accounts—would be to say that only one explanation can be genuinely explanatory. However, this strategy leaves unclear how other arguments—now taken as non-explanatory—could provide insights, which they as a matter of practice do.

For this reason, I support a flexible account of explanation. My deflationary account provides flexibility in distinguishing multiple explanations based on the insights they provide. Adopting my minimal account of scientific explanation, both the commutator and group theoretic derivations of selection rules are properly seen as explanatory: they both recover selection rule phenomena in accordance with physical and mathematical constraints. Having classified these derivations as explanatory, we are then in a position to analyze how they provide

different insights concerning selection rules. I refer to these differences in insights as differences in understanding. The next chapter pursues this account of understanding in earnest by examining these derivations in more detail. Before turning to this analysis, I respond to two alternative accounts of scientific explanation that apply to my case study.

3.2 RELEVANCE ACCOUNTS OF EXPLANATION

In his work on the philosophy of applied mathematics, Chris Pincock has provided a compelling account of mathematical explanations in science. Pincock argues that applied mathematics furnishes explanations by distinguishing relevant features of phenomena from irrelevant features. This is a species of what Batterman and Rice (unpublished) refer to as a common features account. According to Batterman and Rice, common features accounts claim that "a model explains just when it has certain relevant features in common with actual systems and that having these features in common is exactly what does the explaining" (unpublished, p. 1). This view embodies a criterion for explanation: an argument is explanatory only if it references those features that are relevant for a phenomenon of interest while eliminating irrelevant details. In the following sections, I am primarily concerned with the relationship between relevance and explanation. Hence, I refer to positions that adopt this criterion as relevance accounts of explanation. This section characterizes relevance accounts using examples that Pincock considers. In the next section, I develop two related problems for relevance accounts of explanation. These problems help me defend my deflationary account of explanation.

In much of his work, Pincock has focused on accounting for the *content* of mathematical statements employed in science. This account seeks to specify truth conditions for scientific statements involving both physical and mathematical concepts, referred to as "mixed statements." Pincock believes that to characterize truth conditions of mixed statements, we must provide appropriate mapping relations between physical models and mathematical structures. Through these mapping relations, we can specify what must be the case physically for mixed statements to be true. Recognizing a distinction between description and explanation, Pincock admits that an account of content does not automatically yield an account of explanation: "We

should not expect an account of how mathematics describes a target system to be able to provide a complete account of how mathematics can be used to explain features of the target system" (2011, p. 212). Nevertheless, Pincock believes that mappings lead to a natural account of mathematical explanation, where these maps "ground some ways in which mathematics can help in explanation" (2011, p. 213).

Pincock's remarks concerning mapping and explanation illustrate what I call relevance accounts of explanation. Discussing Euler's solution to the bridges of Königsberg problem, Pincock claims that "the explanatory power is tied to the simple way in which the model abstracts from the irrelevant details of the target system. It throws out what is irrelevant and highlights what is relevant" (2011, p. 213). On this view, the argument provides an explanation of the non-Eulerian character of the bridge system by focusing on the abstract graph structure. Focusing on the graph structure eliminates all mention of the bridge system's micro-physical details, material construction, variety of bridge types, etc. Since these details do not impact whether or not the bridge system is Eulerian, they are irrelevant to the explanation. Pincock claims that "what is relevant is the mathematical structure found in the target system itself," providing an example of how his mapping account can extend to an account of explanation (2011, p. 213). A second example further illustrates Pincock's commitment to a relevance account. Discussing an explanation of wave dispersion that relies on two mathematical models—models A and B—Pincock claims that "what we need is the mathematical link between *A* and *B* *because only this allows us to appreciate the relevant features* of the target system" (2011, p. 215, emphasis added). Thus, Pincock claims that relevance is a necessary criterion for explanations in applied mathematics.

If relevance accounts are correct, then my account of explanation is deficient. According to relevance accounts, unless some of the derivations of selection rules latch onto the relevant features of selection rule phenomena, these derivations are not explanatory. To meet the criteria laid down by relevance accounts, an explanation of a selection rule should show how it references relevant physical and mathematical features. This is a stricter criterion than the two minimal constraints I place on explanation in Section 3.1. In the next section, I present two problems for relevance accounts that my alternative account of explanation avoids.

3.3 TWO PROBLEMS FOR RELEVANCE ACCOUNTS

By definition, relevance accounts of explanation rely on a notion of relevance to determine whether or not an argument is explanatory. Once an argument has successfully recovered a phenomenon of interest, relevance accounts instruct us to consider whether or not this argument relies on relevant features while eliminating irrelevant details. Relevance plays a crucial role in grounding the explanatory character of an argument. This reliance on relevance invites a further question: how do we determine which physical and mathematical features are relevant for explaining a given phenomenon? In this section, I argue that in the context of mathematical explanations we sometimes justify ascriptions of relevance by relying on my deflationary notion of explanation. This leads immediately to a circularity problem for relevance accounts: relevance accounts use relevance to ground explanations while at the same time grounding ascriptions of relevance via successful explanations.

In Section 3.3.2, I strengthen this criticism by considering how relevance accounts treat cases of multiple compatible explanations, such as multiple explanations of selection rules. Relevance accounts assume that there is a unique set of relevant features that explain a phenomenon. In cases of multiple compatible explanations, this requires us to determine which explanations invoke this unique set of relevant features. If some multiple explanations rely on different physical features, then we need to determine which features are truly relevant. I argue that in making these determinations, we will again rely on my deflated notion of explanation. This further supports my contention that in some cases of scientific explanation, relevance accounts are circular.

3.3.1 Circularity

According to relevance accounts of explanation, an argument is explanatory when it draws upon the relevant features of a phenomenon. In order for relevance accounts to work, we need some method for distinguishing relevant features from irrelevant features. Without an appropriate method, we will not be able to say why one argument uses relevant features (making it explanatory) while another uses irrelevant features (preventing it from being explanatory). In

general, we do possess a suitable method. This method consists in developing scientific explanations—in my deflationary sense of explanation. By developing explanations of physical phenomena, scientists determine what features are relevant and what features are irrelevant. Unfortunately for relevance accounts, this role for explanations creates a circularity problem: if explanations justify our determination of relevance, then relevance cannot justify explanations. It seems that relevance accounts put the cart before the horse, using relevance to drag along explanations, while in fact it is explanations that guide relevance.

Batterman and Rice (unpublished) raise a similar worry for relevance accounts. They consider an explanation of the universality of some flow patterns present in laminar flow of incompressible fluids. For instance, in the laminar flow regime, incompressible fluids exhibit a parabolic velocity profile. This behavior is universal in the sense that many micro-structurally distinct fluids display this same macroscopic behavior. One explanation of this behavior involves a computational model known as a lattice gas automaton (LGA). In recovering a parabolic velocity profile, this explanation shows that this universal behavior is governed by three fundamental physical properties of fluids: locality, conservation, and symmetry. However, as Batterman and Rice contend, "Simply to cite Locality, Conservation, and Symmetry as being *explanatorily* relevant actually raises the question of why those features *are* the common features among fluids" (unpublished, p. 13). Batterman and Rice argue that in order to explain this universal behavior, we need an account of "*why* those features are common and relevant" (unpublished, p. 22):

> Common features accounts would likely cite the fact that the different fluids have Locality, Conservation, and Symmetry in common as explanatorily relevant and maybe even as explanatorily sufficient. However, as we emphasized in section 3.3 this is a mistake. The fact that the different fluids all possess these common features is also something that requires explanation. (Batterman and Rice unpublished, p. 27)

In Section 3.4, I return to Batterman and Rice's methodological prescriptions for explanations. For now, I argue that recovering the phenomenon of interest provides sufficient grounds for ascriptions of relevancy. This leads me in Section 3.5 to reinterpret Batterman and Rice's prescription as a constraint on good practice, rather than a constraint on genuine explanations.

To see how we determine what features are relevant for explaining a phenomenon, consider the derivations of selection rules examined in Chapter 2. If we adopt a relevance

account, then these derivations are explanatory only if they draw upon the relevant physical details and neglect irrelevant ones. To explain the selection rule on the angular momentum of a transitioning electron, we must know what features of atomic systems bear on this selection rule. Based on the group theoretic derivation, one natural response is to say that the approximate spherical symmetry of atomic systems explains the selection rule on angular momentum. In line with relevance accounts, this means that spherical symmetry is a relevant physical feature of atomic systems for explaining this selection rule. Other physical details—such as the specific atomic number, number of electrons, method of excitation, etc.—are irrelevant for explaining this selection rule.

However, we must ask: how do we determine that spherical symmetry is relevant for this explanation? Due to the fact that conservation of angular momentum arises from the rotational invariance of space, we have a hunch—before undertaking the derivation—that spherical symmetry should matter for explaining a selection rule on angular momentum. Yet, we do not know that spherical symmetry is relevant for this selection rule until we complete a successful derivation that recovers the selection rule. It is only after we see a derivation of this rule from spherical symmetry that we have grounds for saying that spherical symmetry matters in this case. The success of the derivation justifies our ascription of relevance.

Motivated by this example, I argue that a descriptively accurate account of some scientific explanations should invert the relevance relation. It is only after an argument succeeds in recovering an explanandum that we believe particular features we referenced are relevant. If this is correct, then it is inappropriate to use relevance as a criterion for explanatory power: explanatory power is in fact a criterion for relevance. To take relevance as grounding explanation, we would need an account of relevance independent of explanatory considerations. Without such an account, our identification of genuine explanations would be circular. We would designate an argument as explanatory if it drew upon relevant features and eliminated irrelevant details. Yet, at the same time, we would use the explanation to determine what the relevant and irrelevant features were. In other words, we would first assume that we have an explanation to determine relevance. Next, we would rely on relevance to justify that we have an explanation. This chicken and egg scenario points to a problem with the relevance account. We either need to distinguish relevant from irrelevant features independently of any explanation, or else we must reject relevance accounts of explanation.

3.3.2 Multiple Compatible Explanations

Cases of multiple compatible explanations amplify the problem of determining relevance, introduced in the previous subsection. Multiple compatible explanations occur when we have more than one legitimate way of explaining a phenomenon of interest. Derivations of selection rules from both the commutator and group theoretic approaches provide an example. From the standpoint of the commutator approach, the commutation relations are highly relevant for explaining selection rules: by starting with these relations and reasoning appropriately we can derive selection rules. Yet from the group theoretic perspective, these commutation relations do not appear to be relevant: we do not need to use them to derive selection rules. In this section, I argue that multiple compatible explanations pose a problem for relevance accounts.

At least implicitly, relevance accounts seem committed to the idea that there is a unique set of relevant features that explain a given physical phenomenon. A genuine explanation of a physical phenomenon should reference these relevant features and avoid referencing irrelevant features. Cases of compatible explanations pose a problem for this view of explanation. If two compatible explanations rely on different sets of relevant features, how do we determine which set of features is truly relevant? I first illustrate this problem abstractly before returning to my case study. Assume that we have two arguments, A and B, that explain—in my deflationary sense—a phenomenon P. If we adopt a relevance account, our task is to determine whether A or B—if either—*genuinely explains* P by referencing the unique set of relevant features. According to a relevance account, if argument A explains P, then it references all and only the relevant physical and mathematical features that account for P. An analogous statement holds for argument B. Call the collection of physical and mathematical constraints that are relevant to explaining phenomenon P the set R. By definition, a genuinely explanatory argument for P must reference R and no irrelevant statements outside R. Thus, if arguments A and B utilize different sets of physical and mathematical constraints, then only one of them can be genuinely explanatory. Yet, determining which set of features is truly relevant—as opposed to merely apparently relevant—appears to be difficult. Our grounds for claiming that any of these features are relevant stems from their success in explanations—in my weak sense of explanation. Hence, it appears difficult or impossible to determine which set of features is truly relevant since—in the context of compatible explanations—each set of features provides a successful explanation.

By legislating that there is at most a single explanation for a given physical phenomenon, relevance accounts fail to do justice to scientific practice. We frequently have multiple ways of analyzing and understanding a physical phenomenon, and these various points of view stem from multiple explanations. Legislating that there is a unique set of relevant features governing a phenomenon would imply that many of our purported explanations involve irrelevant details. Returning to the present case study, we can examine how this tension might play out in practice. A group theoretic derivation shows that selection rules are the result of symmetries of atomic systems. For instance, the selection rule for the orbital angular momentum quantum number is "due to symmetry under rotation" (Petrashen and Trifonov 2009, p. 267). The rotational symmetry of the atomic system provides a constraint on how the orbital angular momentum can change as a result of electromagnetic radiation. We discover this constraint as a result of applying a general group theoretic criterion for selection rules to the case of rotational symmetry. In deriving this selection rule, we learn that rotational symmetry is relevant while other factors such as the precise atomic weight, number of atoms in an atomic system, experimental apparatus, etc. are irrelevant for the explanation of this selection rule. However, it would be a mistake to say that the commutation relations employed in the commutator derivation are likewise shown to be irrelevant; these relations provide another method for deriving this selection rule. Yet, if we adopt a relevance account of explanation, then only one derivation can employ the unique set of relevant features. Thus, if commutation relations are the relevant features that explain these selection rules, then the group theoretic explanation in terms of irreducible representations is irrelevant, and vice versa. This account of explanation fails to describe the situation we are faced with.

One possible way to diffuse this tension with scientific practice would be to adopt a modified relevance account. We could amend relevance accounts to allow for multiple sets of relevant features that explain a phenomenon of interest. Provided that these sets of relevant features are mutually consistent with each other and—in some sense—embody the "same" information, they can both be taken as playing an explanatory role. Nevertheless, this modified relevance account does not avoid the previous circularity problem. Even if multiple sets of features are taken as relevant, we still need a method—independent of my deflationary sense of explanation—for determining what features are relevant. Once again, I argue that in some cases

of mathematical explanations, our ascriptions of relevance stem from successful explanations, rather than the other way around.

3.4 BATTERMAN AND RICE: MINIMAL MODEL EXPLANATIONS

Batterman and Rice provide an account of a specific kind of explanation that at first glance appears to solve the problems I raise in Section 3.3 for relevance accounts. Building on earlier work by Batterman (2002, 2010), Batterman and Rice develop an account of explanations of universality behaviors. As explained in Section 2.5, selection rules are an example of a universality behavior. Many different kinds of atomic systems exhibit the same selection rules. An important explanatory task is to explain why these micro-structurally distinct systems display macroscopic emission lines governed by the same set of selection rules. To characterize this sort of explanatory task, Batterman and Rice develop an account of "minimal model explanations." In this section, I introduce Batterman and Rice's account. I explain why one might be tempted to see their account as refuting the problems developed in Section 3.3. I then explain why this interpretation is in fact mistaken: the problems of Section 3.3 are actually at the heart of Batterman and Rice's own account.

To provide an account of explanations of universality, Batterman and Rice appeal to minimal models. A minimal model provides a bare-bones description abstracted from the various systems it describes. I use "description" here for lack of a better term. Quoting Nigel Goldenfeld, Batterman and Rice characterize a minimal model as "that model which most economically caricatures the essential physics" (qtd. in unpublished, p. 9). A minimal model fails to accurately describe features of real physical systems. In the case of fluids, an appropriate minimal model allows us to "explore patterns of fluid behavior, while virtually ignoring any realistic details of any actual fluid" (Batterman and Rice unpublished, p. 10). Since these fluid systems exhibit the same behavior, they fall within a universality class. A minimal model explains a universal behavior when (1) it lies within the same universality class as the systems it purports to explain and (2) under perturbations of physical details, these systems and the minimal model flow (in a

topological sense) to the same fixed point. Batterman and Rice illustrate these general remarks with an example from fluid mechanics (unpublished, p. 14-15).

Certain remarks made by Batterman and Rice may be taken as showing that we can—and even must—determine relevance *before* providing an explanation. On this interpretation, Batterman and Rice have marked out asymptotic reasoning and renormalization group techniques as two strategies for determining the relevance and irrelevance of classes of parameters independently of explanations. According to Batterman and Rice, "We require a process for discovering (or demonstrating) why certain dominant features are relevant and why the various heterogeneous features ignored or misrepresented by the minimal model are irrelevant" (unpublished, p. 22-23). Specifically, explanations of universality involve a "process of delimiting a universality class by demonstrating that the details that distinguish the model system and various real systems are irrelevant" (unpublished, p. 2). In cases where we can make determinations of relevance and irrelevance independently of explanatory considerations, my objections against relevance accounts dissolve. In these cases we can determine relevance before recovering a phenomenon to be explained. Hence, we can appeal to relevance in an explanatory capacity.

Indeed, on one interpretation—an interpretation that I shortly contest—this is precisely what Batterman and Rice say about minimal model explanations of universality. This interpretation is supported by remarks such as the following: "By showing that all the other details are irrelevant we can see why only these common features are necessary for the phenomenon to occur" (unpublished, p. 23). Furthermore, Batterman and Rice argue that determining what details are irrelevant plays an essential role in explanation: "There are a number of techniques for demonstrating that a large class of details of particular systems is irrelevant to their macroscale behavior. This is an essential part of the process of delimiting the universality class" (unpublished, p. 24). Out of context, these statements are akin to the relevance accounts discussed in Section 3.1. Similar remarks in earlier writings also ground this interpretation of Batterman. Pincock motivates his own relevance account—in terms of mappings—by attributing a relevance account to Batterman: "The proposal for explanation suggested by Batterman turns on removing irrelevant details and highlighting relevant factors for the phenomenon which is being explained. I have tried to indicate how this account of explanation can benefit from the mapping account of content" (2011, p. 216).

However, a closer reading of Batterman and Rice shows that they do not support a relevance account of explanation. On their account, arguments become explanatory when they answer specific why questions concerning universality. These questions include (1) why a universal behavior requires systems to exhibit certain common features, (2) why these common features are shared, and (3) why differences between these systems are irrelevant for the universal behavior (Batterman and Rice unpublished, p. 13). In providing an explanation of a universal behavior, we learn why certain features are relevant and other features irrelevant. That these features are common and why they are common are results of the explanation, rather than part of the explanation. As Batterman and Rice contend, common features accounts are mistaken because they "[allow] (a version of) the explanandum to masquerade as the explanans" (unpublished, p. 17). On this point, I am in agreement with Batterman and Rice. Nevertheless, I disagree with the restrictions that Batterman and Rice place on genuine explanations. In the next section, I argue that these restrictions should be interpreted as constraints on good practice rather than necessary requirements.

3.5 REINTERPRETING BATTERMAN AND RICE

Both relevance accounts and Batterman and Rice's account of explanations of universality place numerous restrictions on explanations. Relevance accounts require that explanations reference relevant phenomena while not mentioning irrelevant phenomena. As the arguments in the prior two sections show, this restriction on explanation is incoherent. Relevance does not always ground explanations because sometimes explanations ground relevance. In place of relevance, Batterman and Rice constrain explanations of universality by positing specific questions that must be answered. Answering these questions requires some form of stability analysis, such as through asymptotic reasoning or renormalization group techniques. In this section, I argue that Batterman and Rice's account sometimes conflicts with scientific practice. Scientists do not always explain universal phenomena in the way that Batterman and Rice prescribe. To resolve this dilemma, I propose a weaker interpretation of Batterman and Rice's account.

While discussing explanations of universality phenomena, Batterman proposes a strong criterion for this class of explanations. According to Batterman, to explain a repeated pattern of behavior—such as a universality phenomenon—we need to consider the stability of the behavior under perturbation:

> *We need, that is, to ask for an explanation of those very regularities and invariances*. This is the fundamental explanatory question. . . . The answer to this fundamental question necessarily will involve a demonstration of the stability of the phenomenon or pattern under changes in various details. (Batterman 2010, p. 21)

Similarly, Batterman and Rice catalog a number of questions they deem necessary for explaining a universal behavior. These include

(1) Why are a given set of common features necessary for the phenomenon to occur?
(2) Why are non-shared features between systems in the same universality class irrelevant for the phenomenon to occur?
(3) Why do different physical systems have these features in common? (Batterman and Rice unpublished, pp. 13 and 23).

They claim that answering these questions requires some form of stability analysis. First, a minimal model must be found that recovers the phenomenon of interest. Next, the universality class must be delimited by showing how the systems of interest and this minimal model all flow to the same fixed point under perturbations of unshared details. As Batterman and Rice claim, "The real explanatory work is done by showing *why* the various heterogeneous details of these systems are irrelevant and, *along the way, by demonstrating the relevance of the common features*" (unpublished, p. 17). Explanations of universality require an argument that answers these three questions.

However, scientists do not always provide answers to these questions when explaining universal behaviors, as witnessed by the derivations of selection rules presented in Chapter 2. As I argued in Section 2.5, selection rules are properly seen as an example of universal behavior. Hence, explanations of selection rules fall within the scope of Batterman and Rice's account of explanations of universality. To satisfy Batterman and Rice's explanatory criteria in the case of selection rules, we would need to derive selection rules in a more restrictive manner than the

derivations presented in Chapter 2. As in those derivations, we would first have to provide an argument that recovers selection rules. However, we would also have to show why these selection rules are stable under changes in atomic number, ionization, excitation source, etc. We would have to show that even when we perturb an atomic system from idealized spherical symmetry, the same selection rule phenomena still dominate behavior. One way to do this would be to treat a model of a hydrogenic atom as a minimal model for this class of systems. We might then be able to show that as we perturb the details of many-electron atoms, these systems and this minimal model "flow" toward a shared fixed point. Physicists could then use this analysis to answer variants of the three aforementioned questions that Batterman and Rice pose. However, in practice, physicists have not answered these questions, at least not in most presentations of selection rules.

Demonstrating the stability of selection rules under perturbation in atomic details would further justify the derivations of selection rules presented in Chapter 2. Nonetheless, this kind of stability analysis should not be a requirement for explanation. In practice, we learn that these selection rules are stable under perturbation by the experimental fact that a wide variety of atomic systems exhibit them in their spectra. The commutator and group theoretic derivations show that these selection rules follow from accepted mathematical and physical constraints. Although these derivations are given for an idealized spherical system, its success gives us reason to believe that the features identified—e.g. commutation relations or irreducible representations—are relevant to selection rules in real atomic systems. Quantifying the extent to which these features impact real atomic systems would further justify why these derivations work, but it is unnecessarily restrictive to make this a requirement for explanation. This requirement would mean that the derivations in Chapter 2 are mere derivations and not explanatory. It would also mean that most—if not all—texts on quantum mechanics derive selection rules but fail to explain them. Hence, I take this case study to support the claim that Batterman and Rice's constraints on explanation are overly restrictive.

A weaker interpretation of Batterman and Rice's account of explanation dissolves this tension with scientific practice. Rather than interpret their demand for stability analysis as a strict requirement for explanations of universality, I propose that this demand should be taken as a constraint on good explanatory practice. On this view, Batterman and Rice's criteria are suggested guidelines that apply when scientists want to answer a particular kind of why-question.

In the context of selection rules, this question would go something like the following, "What justifies the applicability of the central field approximation to elements with atomic numbers less than 40?" Batterman and Rice arguably believe that explaining this requires a stability analysis. However, on a weaker interpretation, an explicit stability analysis is not necessary for an explanation of selection rules. Instead, we can view the applicability of the central field approximation as lying in the background of the derivations presented above. This recognizes that scientists' explanations are often elliptical in nature. Actual scientific explanations do not always provide detailed justification for the assumptions they employ. This descriptive claim provides my main reason for disagreeing with the stronger interpretation of Batterman and Rice considered above. I have no such quarrels with the weaker interpretation.

Additional explanatory criteria could certainly be added to the two minimum criteria of my deflationary account of explanation. However, I contend that any additional criteria should be viewed as methods to distinguish the *quality* of explanations rather than to determine whether or not an argument is explanatory *per se*. This points to a key difference between my preferred way of talking and Batterman and Rice's. Whereas Batterman and Rice prefer to view these additional constraints as restrictions on explanation, I prefer to view them as methodological constraints on understanding. I believe that my preferred way of talking opens the door to a more fruitful analysis of how mathematics contributes to scientific understanding. Rather than debate whether a particular mathematical argument is explanatory, we can focus on how mathematical ideas within an argument provide physical insights. This motivates an analysis of a wider variety of argument strategies. In some cases, this way of talking sits well with Batterman's own remarks on scientific understanding:

> Nevertheless, the asymptotic investigation of this equation is essential for an understanding of why rainbows always appear with the same patterns of intensities and spacings of their bows. The asymptotic investigation of the wave equation leads to an understanding of the *stability* of those phenomena under perturbation of the shape of raindrops and other features. (Batterman 2010, p. 21)

On my view, asymptotic reasoning can play an important role in understanding without being a prerequisite for explanations of certain phenomena.

3.6 SUMMARY

I began in Section 3.1 by presenting a deflationary account of mathematical explanations. On my account, an argument is explanatory if it recovers the phenomenon of interest according to principled mathematical or physical constraints. I have argued that this minimal notion of explanation should be taken as basic: it lies at the root of philosophers' richer notions of genuine explanation. By separating our usual notion of "genuine explanation" into a minimal sense of explanation and a further sense of scientific understanding, I believe that my preferred way of speaking avoids the problems I develop for alternative accounts of explanation. In particular, relevance accounts of explanation seem saddled with the problem of determining a unique set of physical and mathematical features relevant to a given explanandum. I have argued that we sometimes determine what features are relevant only by constructing successful arguments that meet my minimal criteria for explanation. This poses a circularity problem for relevance accounts. Furthermore, relevance accounts have difficulty acknowledging multiple compatible arguments as genuinely explanatory. This creates tension with scientific practice.

Similarly, Batterman and Rice's account of explanations of universality behaviors places restrictions on explanation that scientists do not always honor. For instance, derivations of selection rules generally do not satisfy Batterman and Rice's criteria, even though selection rules are an example of universality. Hence, I have argued for a weaker interpretation of Batterman and Rice's view, wherein their criteria are seen as prescriptions for good explanatory practice. In the next chapter, I consider features of selection rule derivations that are separate from explanation. These features—such as the linguistic resources and organizational structure that we use to derive selection rules—are relevant for an analysis of how the commutator and group theoretic approaches provide understanding. I rely on expressive and structural differences between the two approaches to articulate how group theory provides a superior understanding of selection rules.

4.0 UNDERSTANDING SELECTION RULES IN ATOMIC SPECTROSCOPY

Chapter 3 sought to articulate how the commutator and group theoretic derivations of selection rules are explanatory. I argued for a minimal account of explanation, wherein arguments are explanatory when they show how a phenomenon follows from accepted mathematical and physical constraints. This bare view of explanation alone does not distinguish the intellectual content of the commutator and group theoretic derivations. This chapter accommodates these differences by developing an account of scientific understanding. Following Manders (unpublished), I reserve "understanding" for a body of knowledge shared amongst competent practitioners. As such, understanding has little to do with the idiosyncrasies of particular scientists. Rather, understanding operates at a level of shared intellectual content. With this picture of understanding in mind, I argue that scientific understanding consists in an organizational structuring of physical and mathematical constraints.

To support this claim, I examine the similarities and differences of the commutator and group theoretic derivations of selection rules. This philosophical analysis would be enriched by considering multiple phenomena in atomic spectroscopy, such as the Zeeman and Stark effects, the coupling of angular momenta, the characterization of the eigenvalues of the hydrogen spectra, and the Wigner-Eckart theorem. Nevertheless, I focus on selection rules in order to show that alternative mathematical frameworks can impact the understanding of a specific problem. Differences between the frameworks point to differences in the understanding that the derivations provide.

In Section 4.1, I distinguish two senses of "better explanation." A traditional conception of theoretical virtues captures the first sense of better explanation. This sense focuses on the ways in which some explanations might be closer to the truth than others. It is motivated primarily by debates concerning how data underdetermines theory. However, in my case study, the legitimacy of the group theoretic and commutator approaches is not in question: both

approaches provide valid explanations of selection rules. Thus, an analysis of approximate-truth is not relevant for an analysis of differences in understanding. This motivates a second sense of "better explanation." This sense focuses on how mathematical and scientific concepts provide more illuminating and intelligible explanations of phenomena. Explanations that are better in this regard are better *qua* understanding. In subsequent sections, I argue that this intellectual advantage consists in organizing information more effectively.

To develop an account of understanding, I begin with an analysis of expressive means. In the sense used by Manders, expressive means comprise the conceptual resources available for thought within an intellectual framework. In Section 4.2.1, I catalog mathematical concepts employed by the group theoretic and commutator approaches. This leads to an analysis of how the approaches differ in mathematical resources. Section 4.2.2 articulates a notion of expressive power distinct from expressive means. Expressive power describes the range of physical phenomena that a framework can discuss. Some physical features are only intellectually accessible under suitable expressive means. Nonetheless, two frameworks can employ alternative mathematical concepts and still talk about identical physical features. I show how differences in understanding are partially constituted by differences in expressive power—differences that result from alternative expressive means.

Supported by these considerations of expressive means and expressive power, Section 4.3 constructs a positive view of understanding. Each expressive means yields a particular structuring of physical knowledge. These structures differ in how they organize knowledge, accounting for subsequent differences in theoretical virtues such as modularization and tractability. In virtue of their differences in expressive means (and subsequently expressive power), the commutator and group theoretic approaches present selection rules differently. I contend that understanding is constituted by an organizational structure and its concomitant theoretical virtues. By analyzing how organizational structures can be intellectually advantageous, we come to see how the group theoretic approach provides a deeper understanding of selection rule phenomena. This analysis constitutes an explication of Wigner's belief that group theory provides a natural language for expressing selection rules (1959, p. v).

4.1 TWO SENSES OF SUPER-EMPIRICAL VIRTUE

When group theory is applied to quantum mechanics, there is a shared sense that this application is of tremendous intellectual power. This raises two questions: What does this enhanced intellectual power consist in? How does group theory provide this heightened understanding? We begin to answer these questions by recognizing that group theory provides explanations of quantum mechanical phenomena. However, an appeal to explanations provides only part of an answer. As we have seen in the specific case of selection rules, many aspects of atomic spectra do not require group theoretical arguments for their explanation. The commutator approach provides explanations too. Thus, if we are going to account for how group theory strengthens our grasp of these aspects and what this strengthening consists in, it is insufficient to point to the fact that group theory provides explanations. We must look to the *way* in which these explanations are given. In certain cases, group theoretic explanations are superior. We can articulate how group theory presents an intellectual advance by fleshing out how group theory provides better explanations. This requires that we first consider what it means for an explanation to be better.

Standard accounts of inference to the best explanation (IBE) contend that better explanations are closer to the truth than competing explanations. Proximate truth is in turn measured by appealing to super-empirical virtues. These theoretical virtues—such as simplicity, ontological parsimony, fruitfulness, and generalizability—are taken as marks of truth. Other factors being equal, the more theoretically virtuous explanation is assumed to be closer to the truth and hence better than less virtuous explanations. Scientific realists developed this sense of "better explanation" to rebut underdetermination challenges from scientific antirealists. In the context of those debates, we are challenged to distinguish two competing explanations of a phenomenon. If both explanations are empirically adequate, IBE suggests that we catalogue their respective super-empirical virtues to determine which one is more likely to be true or, alternatively, closer to the truth.

This IBE prescription for distinguishing explanations is inappropriate for distinguishing the group theoretic and commutator approaches. As discussed in Chapter 1, these approaches are compatible rather than competing. Hence, the group theoretic and commutator explanations of selection rules do not fall within the context of an underdetermination challenge. Since we do not need to decide which approach to accept, we do not need to consider how their differences

impact proximate truth. Yet, according to the IBE account of "better explanation," explanations provide more understanding when they are closer to the truth. IBE has little more to say about understanding beyond this relationship to proximate truth. Consequently, this first sense of "better explanation" does not bring us closer to an account of how the group theoretic and commutator approaches differ in the understanding they provide.

Cases of multiple compatible explanations point to a different sense of "better explanation." We need to distinguish compatible explanations, but we cannot appeal to proximate truth. Since these explanations differ in the insights they provide, we need an account of "better explanation" *qua* understanding, rather than *qua* proximate truth. This second sense of better explanation focuses on how an explanation can provide a superior understanding of a problem, even when alternative explanations are just as likely to be true. Characterizing this second sense is difficult insofar as analyzing scientific understanding is difficult. Although this analysis does not depend on truth-likeliness, the methodological focus of IBE on virtues remains relevant.

To articulate intellectual differences between compatible explanations, I propose to focus on super-empirical virtues. My strategy differs from IBE because these virtues are not used as indicators of truth-likeliness. In Section 4.3, I distinguish the group theoretic and commutator explanations based on modularization, tractability, uniformity of treatment, and unification. With the exception of unification, these are nontraditional super-empirical virtues. They are relevant for articulating differences in understanding, even though they may have little to do with tracking truth. These virtues focus attention on how different approaches organize information. Alternative frameworks lead to different solution procedures for solving problems. These procedures differ in how they decompose a problem into sub-problems (modularization), and how tractable these sub-problems are. As pointed out above, these differences in organizational virtues stem from differences in the mathematical and physical language used to express problems. Hence, I begin by describing expressive differences between the group theoretic and commutator approaches.

4.2 EXPRESSIVE MEANS AND EXPRESSIVE POWER

Before articulating the sense in which one understanding of a phenomenon can be better than another, we must consider the sense in which two ways of understanding can be different. In full generality, this issue is too expansive to be treated here. Instead, I focus on how the group theoretic and commutator approaches provide different understandings of selection rules. Comparing the two approaches, a salient difference is the language used to handle relevant features of the phenomena. Although there are points of overlap, the two approaches employ different concepts. These different concepts arise in the context of implementing the same solution strategy. Perturbation theory tells us that transition intensities are proportional to the square of the perturbation operator's matrix elements. From this fact, it is clear that we should find the conditions under which these matrix elements necessarily equal zero, in order to eliminate unnecessary computations. Both the group theoretic and commutator approaches implement this strategy, but they do so in strikingly different ways. They exhibit differences not only in expressive means—the language used to discuss the phenomenon of interest—but also in expressive power—the features of phenomena that can be talked about. This latter difference in expressive power is itself a consequence of differences in expressive means. Ultimately, these differences in expressive means and expressive power result in important organizational differences that affect the understanding of selection rules. In Section 4.3, I argue that the organizational structures of the two approaches should be identified with the understanding they provide.

4.2.1 Characterization of Expressive Means

As a working definition of expressive means, I adopt the following:

> The expressive means of a given conceptual framework are constituted by the mathematical and physical concepts that it uses.

These concepts are brought to bear on issues of description, prediction, and explanation. They are the soil from which understanding sprouts. To support this claim, I show how salient

intellectual differences between the group theoretic and commutator approaches result from differences in expressive means. Although the two approaches have much in common, their differences cause a nontrivial change in understanding. After laying out concepts shared by both approaches, I discuss key differences. This paves the way for an analysis of expressive power in Section 4.2.2.

The group theoretic and commutator approaches have considerable overlap in expressive means because they both utilize a Hilbert space formalism for quantum mechanics. Unsurprisingly, this commonality yields considerable overlap in expression. Hermitian operators are associated with physical observables, where possible measurement outcomes correspond to eigenvalues of these operators. Eigenvectors of the Hamiltonian operator correspond to stationary states of the system. More generally, eigenvectors of Hermitian operators correspond to physical states that have well-defined values of the corresponding observable. Both approaches assume that quantum states can be characterized by a complete set of commuting observables, $\{T\}$. In both approaches, we explicitly work with the angular momentum basis where J^2 and J_z are diagonal, enabling us to label the eigenvectors of the Hamiltonian with the quantum numbers j and m, where j indexes the total angular momentum of the stationary state and m indexes the z-component of the angular momentum. We denote the other members of the set $\{T\}$ collectively by $\{A\}$, with eigenvalues $\{\alpha\}$. This enables us to label eigenkets by these eigenvalues, denoting an arbitrary stationary state by $|\alpha jm\rangle$.

Distinctively, the commutator approach relies on commutation relations between operators. The orbital angular momentum operator is defined as the cross product of the position and momentum operators. Using this definition and the canonical commutation relations for position and momentum, we derive the cyclic commutation relations for orbital angular momentum. Assuming that analogous relations also hold for the intrinsic spin angular momentum, we derive the general commutation relations for angular momentum. From these, the other relevant commutation relations follow, enabling us to derive selection rules.

Commutation relations do not form any part of the expressive means of the group theoretic approach. Since the concept of commutation relations follows directly from a Hilbert space formalism, it is not that the group theoretic approach is conceptually barred from using commutators. Rather, the group theoretic approach to selection rules does not need to express commutation relations. Instead, we take what could rightly be seen as detours through group

theory and group representation theory. As we will see in the next section, this alternative path allows us to express physical information that the commutator approach cannot access. This, in turn, allows us to restructure our approach to finding selection rules.

By applying group theory to quantum mechanics, we develop additional expressive means beyond those employed by the commutator approach. These alternate expressive means allow us to focus on other features of quantum mechanical phenomena—most notably, their symmetries—providing a different way of thinking about selection rules. Moving beyond individual operators that commute with the Hamiltonian, we focus on the set of operators commuting with the Hamiltonian. These operators form a group, corresponding to symmetries of the given quantum mechanical system. We call the group of operators that commute with the Hamiltonian the "group of the Schrödinger equation." In the specific case addressed in this essay, this group is the full orthogonal group O(3), which is the direct product of SO(3) and the reflection group. Through the representation theory of groups, we are able to express unitary irreducible representations of this group and direct products of these representations. We label the eigenkets of the Hamiltonian operator via irreducible representations of the symmetry group, and the perturbation operator as a direct sum of irreducible representations. Thus, the group theoretic approach eschews commutators for talk of irreducible representations. This difference in expressive means accounts for key differences in expressive power.

4.2.2 Differences in Expressive Power

As the Hilbert space formalism for quantum mechanics makes clear, expressive means and expressive power are closely related. In this case, the theory of self-adjoint operators provides suitable expressive means for discussing physical states. The expressive means blends seamlessly into the phenomenon expressed. This indicates that it may not be possible to sharply distinguish expressive means from expressive power. Fortunately, a precise distinction is not necessary for the two notions to be useful. Whereas expressive means are the accepted linguistic resources within a domain of thought, expressive power is the range of phenomena that can be talked about. I intend these as working definitions for use in distinguishing the commutator and group theoretic approaches. This section characterizes the physical features that the two approaches allow us to focus on.

The clearest cases of alterations-in-expressive-means affecting understanding arise from expansions of expressive power. In transitioning from classical mechanics to quantum mechanics, physicists made a new class of phenomena amenable to description, prediction, and explanation. This change in expressive means led to a radical change in expressive power. Such cases of radical change are interesting; it is worth considering in detail how such differences impact scientific understanding. However, compared to the case I consider below, cases of radical changes in expressive power are relatively unproblematic. Quantum mechanics provides a superior understanding of atomic particles because it treats problems that lie outside the scope of classical mechanics. There is more to be said here, but an expansion of expressive power plays a central role in constituting these differences in intellectual understanding. Since this expressive expansion makes certain problems soluble that were not soluble before, concomitant changes in understanding are relatively unmysterious.

There are also uninteresting cases, cases where alternative expressive means share equivalent expressive power and affect few or no differences in understanding. Notational variants exemplify this case. In merely changing symbols we use to describe a phenomenon—substituting sign for sign—we alter neither the phenomena we can talk about nor the problems we can solve. Two notational variants say exactly the same. A similar case arises in linear algebra, where a single vector can be expressed in terms of different bases. If these bases are related by an isometry, then the length of the vector remains invariant. Hence, this change in expressive means (this basis change) preserves an aspect of expressive power (vector length). Even if the given similarity transformation is not given by an isometry, both bases still have sufficient expressive power to characterize vector length. This example takes us a little further from notational variants, for now the vector length is not necessarily preserved. Nevertheless, not much has changed. Any standard problem amenable to treatment in the first basis is most likely just as amenable in the second. These cases are uninteresting because their differences in expressive means have not affected a difference in understanding. This is likewise unsurprising: if there are no differences in expressive power—both in the problems that the expressive means can solve and the features of phenomena they focus on—then altering expressive means does not alter understanding.

Cases lying between these two extremes are philosophically more challenging. There are cases where crucial aspects of expressive power remain invariant while the alternative expressive

means nonetheless underpin nontrivial differences in understanding. In these cases, alternative expressive means can address problems of interest, even as we recognize that an intellectually powerful development has occurred. This makes an analysis of understanding more difficult than in cases such as the transition from classical mechanics to quantum mechanics. As a simple example, consider spherical and rectangular coordinates as two alternative expressive means. Even though elementary problems soluble in spherical coordinates are soluble in rectangular coordinates, the solution procedures are not always equally tractable. Even though a crucial aspect of expressive power remains invariant under these nonlinear coordinate changes—namely, the problems we can solve—coordinate changes are not notational variants. In moving from rectangular to spherical coordinates, we focus on different features of geometry. Rather than attend to the extension of a vector along three orthogonal axes, we identify it with its radial distance and polar and azimuthal angles. Although we have definite methods for transforming between these representations, we have not simply replaced sign for sign. Nonlinear coordinate changes restructure the information characterizing a geometrical situation. These restructurings lead to organizational advantages and disadvantages: problems with spherical symmetry are more tractable in a spherical coordinate system. I return to this notion of organizational advantage in the next section.

The group theoretic and commutator approaches provide an example of this interesting intermediate case. They share aspects of expressive power but are not notational variants. Some overlap in expressive power arises from shared expressive means, such as the Hilbert space formalism. These commonalities are not philosophically problematic. For instance, both the group theoretic and commutator approaches utilize the concept of a matrix element and refer to matrix elements in the same way. This follows from their adoption of the Hilbert space formalism for quantum mechanics. These commonalities in expressive power are trivial: they occur because both approaches share expressive means. However, the derivations of selection rules represent a fundamentally distinct kind of expressive invariant. Even though both approaches state and derive these rules—exhibiting an overlap in expressive power—they do so via *different* expressive means. These varied expressive means lead to differences in expressive power that nonetheless both suffice for explaining the selection rules of interest. Hence, the approaches are not an example of a radical change in expressive means, wherein the problem of interest—explaining selection rules—is accessible from only one vantage point. We face a more

subtle question of how two approaches can adequately solve the same problem while providing different understandings.

A partial answer to this question stems directly from differences in expressive power. We understand selection rules differently from the vantage point of the two approaches because we see selection rules as following from different features of phenomena. Although both approaches begin at the same point—with matrix elements—and end at the same point, i.e. with selection rules, it is not the case that each step within the derivations can be translated from one approach to the other. The group theoretic approach expresses ideas that have no analog in the commutator approach, and vice versa. For example, the group theoretic approach has suitable expressive means to label eigenkets according to their symmetry type (irreducible representations) within the group of the Schrödinger equation. The commutator approach—as it stands—lacks suitable expressive means to communicate this feature of eigenkets. Although commutation relations can express certain symmetry relations between operators, they do not make symmetries of eigenkets available for consideration. For its part, the group theoretic approach is unable to express some of the detailed relationships between operators that commutation relations embody. The closest analog to these relationships is the direct product formed between the (in general sums of) irreducible representations that the operators transform as. From the group theoretic vantage point, we understand selection rules as consequences of restrictions on direct products of irreducible representations, which are symmetry types. From the commutator approach, we understand selection rules as consequences of canonical commutation relations and the restrictions on matrix elements that they provide.

Nevertheless, merely focusing on these differences in expressive power does not satisfactorily characterize how the two approaches provide different understandings of selection rules. From what has been said so far, there is no reason to treat one of these approaches as decisive. There are no grounds for saying that the application of group representation theory to atomic spectroscopy provides an intellectual advance. Somehow, intellectual asymmetries develop between alternative expressive means. These asymmetries arise from differences in expressive power, but characterizing them requires more than marking out expressive differences. These further considerations can be broadly construed as the relative theoretical virtues of alternative expressive means. As argued in Section 4.1, the relevant sense of theoretical virtue is of a wholly different nature than its usual meaning within the scientific

realism/anti-realism debates. I propose to characterize these considerations as organizational advantages. These advantages are not trivial. Their intellectual value extends far beyond presentational niceties.

4.3 UNDERSTANDING AS ORGANIZATIONAL STRUCTURE

My goal in this section is to further unpack our pre-theoretical notion of understanding. In the previous section, I argued that scientific understanding partially consists in the expressive means that a framework adopts and the resulting expressive power. Expressive means and expressive power manifest the understanding that a framework provides. In what follows, I aim to show how scientific understanding relates to the organizational structure of an explanation. I assume that understanding depends partially on the approach used to explain a phenomenon. In other words, I assume that solutions to problems are understood in light of the solution procedure used to obtain them. Even if two approaches yield the same solution to a problem, they can provide different understandings in virtue of using different procedures. Granting this assumption, I explicate our pre-theoretic notion of understanding by analyzing the solution procedures provided by the commutator and group theoretic approaches. These solution procedures have different organizational structures. I argue that a superior organizational structure equates with superior understanding. To characterize how a solution procedure can be superior compared to another, I appeal to theoretical virtues, which I refer to as organizational advantages. These organizational advantages include modularization, effective structuring of a problem's search space, tractability, motivation, and uniformity of treatment.

In Section 4.3.1, I present the physical-mathematical principles that underlie the commutator and group theoretic approaches. I provide outlines of these approaches to facilitate an analysis of their organizational structures. I argue that the group theoretic approach does a better job of breaking derivations of selection rules into sub-problems. This enhanced modularization provides some immediate advantages in terms of identifying difficulties and understanding how variations in parameters affect solutions. In this case, enhanced modularization also facilitates a superior structuring of the search space: whereas a key step of

the commutator approach requires a search procedure, the steps in the group theoretic approach provide a decision procedure. Additionally, the steps in the group theoretic approach are easier to implement, making them more tractable. Finally, I briefly consider how the search-procedure step in the commutator approach makes motivating a suitable commutation relation difficult. The group theoretic approach does not suffer from an analogous difficulty regarding motivation.

In Section 4.3.2, I consider two additional organizational advantages: uniformity of treatment and unification. Although these two theoretical virtues are sometimes equated, I suggest a way of distinguishing them based on an example from my case study. I argue that a scientific framework can provide a uniform treatment of a set of phenomena without thereby providing a unification of these phenomena. According to the terminology I adopt, unifications require an additional condition beyond those required for a uniform treatment. Loosely, I characterize uniform treatments as derivations of phenomena using the same solution procedure. To *unify* two phenomena, an approach must treat both phenomena within a single argument, showing how the phenomena are intertwined as part of a single, unified viewpoint.

4.3.1 Modularization and Tractability

We can appreciate the structural differences between the commutator and group theoretic approaches by juxtaposing the strategies they provide for deriving selection rules. At the heart of each approach is an underlying principle. In the commutator approach, we utilize Dirac's principle (confer Section 2.2.3). This principle provides one method for determining when matrix elements of a perturbation operator are necessarily zero. When a matrix element connecting an initial state to a final state is necessarily zero, a transition cannot occur between those two states under the given radiative mechanism.

Dirac's Principle:

An equation of the form (4.1) specifies a necessary condition for matrix elements of the perturbation operator to be nonzero.

$$\sum_r f_r P g_r = 0 \tag{4.1}$$

Equation (4.1) is an algebraic equation that is linear in the perturbation operator, P. Provided that we have specified stationary states of our atomic system in terms of a complete set of commuting observables A, f_r and g_r depend only on operators within the set A.

In the group theoretic approach, we utilize what I call Wigner's principle (confer Section 2.3.1). This principle also provides a necessary condition for non-vanishing matrix elements of the perturbation operator.

Wigner's principle:

If the triple direct product $\Gamma = (\Gamma^i)^* \otimes \Gamma^p \otimes \Gamma^f$ does not contain the trivial irreducible representation, Γ^0, of $\mathcal{G}$, then the matrix element $\langle \psi_i | P | \psi_f \rangle$ necessarily vanishes.

Under a change of basis, the triple direct product $\Gamma = (\Gamma^i)^* \otimes \Gamma^p \otimes \Gamma^f$ decomposes into a direct sum of irreducible representations. If this decomposition does not contain the trivial irreducible representation, Γ^0 then a transition cannot occur between initial and final states that transform as Γ^i and Γ^f, respectively. Hence, in order for a transition to be allowed between these states, $\Gamma = (\Gamma^i)^* \otimes \Gamma^p \otimes \Gamma^f$ must contain Γ^0.

Dirac's and Wigner's principles illustrate how alternative expressive means can organize solution strategies differently. Implicit in each principle is a series of steps to follow. These steps represent how each principle modularizes the problem of deriving selection rules. An organizational advantage, modularization focuses attention on specific aspects of the concepts referenced in a solution procedure. By focusing on these specific aspects, modularization allows a problem to be decomposed into sub-problems. After presenting the commutator and group theoretic solution strategies, I argue that group theory provides a superior modularization. This modularization enables a more effective structuring of the search space, leading to enhanced tractability.

Dirac's derivation procedure—introduced in Section 2.3.3—consists of three steps:

Commutator Approach [Method C]:

C1.) Determine a complete set of commuting observables.

C2.) Find a commutation relationship linear in the perturbation operator that is a sum of products of the perturbation operator and an operator diagonal in the basis.

C3.) Use the commutation relationship from C2 to compute an algebraic restriction on the matrix elements of the perturbation operator in terms of the eigenvalues of the diagonal operator.

We use the complete set of commuting observables from step C1 to label our basis kets. In the derivations presented in Chapter 2, we focused on simultaneous eigenkets of the square of the angular momentum operator (L^2) and it z-component (L_z), labeled by corresponding eigenvalues l and m: $|lm\rangle$. In the simple case of Section 2.2.1, we notice that the commutation relation between the z-component of the position operator and the operator for the z-component of angular momentum satisfies constraint C2:

$$[L_z, z] = L_z z - z L_z = 0.$$

L_z is diagonal in the standard angular momentum basis, and z is proportional to the z-component of the perturbation operator. Hence, the commutator expression $L_z z - z L_z$ is a sum of terms that are both linear in the perturbation operator (z is raised to only the first power) and products of the perturbation operator and an operator diagonal in the basis. At this point, we know immediately that a selection rule is at hand. C3 tells us that we simply need to compute the selection rule by calculating the expectation value of the commutator $[L_z, z]$. This tells us that the matrix element $\langle l'm'|z|lm\rangle$ equals zero unless $m' = m \Rightarrow \Delta m = 0$, where m represents the z-component of the orbital angular momentum.

Wigner's principle provides an alternative solution procedure, which I refer to as Method G. Notably, it is not the case that Wigner's group theoretic procedure provides a further modularization of Dirac's procedure. The group theoretic approach does not break the steps of Method C into additional components. Rather, it employs alternative expressive means to structure the solution procedure in a way that is completely different. For instance, Method G does not provide additional information to find a suitable commutation relation satisfying step C2.

Group Theoretic Approach [Method G]:

G1.) Determine the group of the Schrödinger equation, $\mathcal{G}$.
G2.) Determine the direct sum of irreducible representations to which the perturbation operator belongs.
G3.) Label the irreducible representation that the initial wavefunction belongs to.

G4.) Compute the direct product of representations of the perturbation operator, initial wavefunction, and final wavefunction.

G5.) Apply Wigner's principle to determine which irreducible representations the final wavefunction can belong to.

G6.) Translate the restriction on irreducible representations to a restriction on eigenvalues.

Method G represents an organizational advantage over Method C in part because Wigner's principle provides a superior modularization. It focuses attention on more detailed aspects of the physical system, allowing us to decompose derivations of selection rules into distinct sub-problems. Implicit in Wigner's principle is the determination of the group of the Hamiltonian, $\mathcal{G}$. This group specifies the symmetry of the given physical system uniquely up to isomorphism. Whereas Dirac's principle instructs us to focus on the basis functions of a complete set of commuting observables, the group theoretic approach narrows our focus to the relevant symmetry group. This narrowed focus constrains the rest of the solution. It tells us that the first thing we need to do is determine the group of the Hamiltonian. This step (G1) is the first sub-problem that Wigner's principle provides. In contrast, the initial step of the commutator approach (step C1) is overspecified: it tells us to label our wavefunctions using a complete set of commuting observables. In practice, we do not actually find an explicit set of commuting operators. For instance, to find selection rules for angular momentum, we focus on the operators L^2 and L_z and work in the standard angular momentum basis.

Turning to each method's second step, both approaches focus attention on the perturbation operator. Yet, the group theoretic approach emphasizes a particular property of the perturbation operator, namely, its symmetry type. Rather than examine the perturbation operator wholesale, group theory tells us to focus on the representation that the perturbation operator transforms as. This focus introduces a sub-problem that can be performed independently of the remaining parts of the derivation: Determine the direct sum of irreducible representations to which the perturbation operator belongs. Step G3 presents a distinct sub-problem: once we have determined the symmetry group of the Hamiltonian, we can label the symmetry types of the final and initial wavefunctions in terms of irreducible representations. Whereas the commutator approach focuses attention on entire wavefunctions, the group theoretic approach focuses attention on the symmetries of wavefunctions. This heightened focus facilitates the introduction

of sub-problem G3. With steps G2 and G3 complete, we can complete the next sub-problem: computing the direct product of the representations. An application of Wigner's principle (G5) and a final interpretation step (G6) provide selection rules.

The enhanced modularization of Method G is organizational in nature because it pertains to how group theory restructures the problem of deriving selection rules into sub-problems. It takes the larger problem of determining necessary conditions for non-vanishing matrix elements of the perturbation operator and directs attention to specific aspects of the atomic system, the perturbation operator, and the initial and final wavefunctions. Through Wigner's principle, group theory specifies precisely how the symmetries of the perturbation and wavefunctions govern selection rules. By decomposing a problem into sub-problems, modularization provides two advantages.

First, if we run into difficulties, modularization helps us pinpoint where those difficulties arise. The group theoretic approach specifies a set of facts that are sufficient for deriving selection rules, including the group of the Schrodinger equation, the representation of the perturbation operator, and the triple direct product decomposition. If we were to get stuck in our derivation, it would be because we encountered difficulties in obtaining one of these pieces of information. Modularization provides an organizational advantage because it disentangles these pieces of information. For instance, if we get stuck finding a representation for the perturbation operator, we know that to solve this problem we do not need to focus more on the symmetry group of the problem (G1) or focus on labeling the symmetries of the wavefunctions (G3). Provided that we have completed step G1, we do not need to tinker with it more to solve our difficulty with step G2. Likewise for step G3. By disentangling this information, we are better able to direct attention to solving sub-problems. In contrast, if we run into difficulties in determining a suitable commutation relation, step C2 provides little guidance about where our problem lies. Step C2 tells us the form our commutation relation must take, but it does not decompose this problem into sub-problems. It does not disentangle how the perturbation operator and other operators relate to one another, or what role the eigenfunctions play. The lack of further modularization creates an organizational disadvantage. The commutator approach fails to organize physical relations as effectively as the group theoretic approach.

A second advantage of modularization is increased awareness about how varying the content of a problem affects the solution. For instance, step G1 makes it clear that the selection

rules of a physical system depend on its symmetry group. If we vary the symmetry group of a problem, we understand that the selection rules might change. Step G2 affords a similar awareness with regards to the perturbation operator. It makes apparent that finding selection rules for a higher order perturbation or an alternative radiative mechanism could alter the symmetry type of the perturbation operator, therefore leading to different selection rules. We also see that certain features of the problem—such as the overall symmetry group and the symmetry types of the wavefunctions—remain invariant under changes in the perturbation operator. Similar remarks pertain to step G5, where we implement Wigner's symmetry principle. Due to the precision of this step in focusing attention on irreducible representations, it is possible to back-calculate the class of symmetry-types necessary for a perturbation to allow a transition between initial and final wavefunctions of given symmetries. In this way, modularization affords flexibility in setting up problems. Not only can we determine selection rules under a given radiative mechanism, but also we can determine what radiative mechanisms would allow a transition between specified initial and final states. Method C is not sufficiently modularized to support an analogous back-calculation procedure.

These differences in modularization directly impact how the two approaches structure the search space when deriving selection rules. The commutator approach modularizes the problem into a search procedure, whereas the group theoretic approach provides a decision procedure for computing selection rules. Like modularization, this difference is organizational: it deals with how we have related physical and mathematical constraints to construct a solution procedure. Consider the form of Method C first. Step C2 delimits a class of suitable commutation relations, namely, those that are sums of products of diagonal operators where each term is linear in the perturbation operator. However, step C2 does not specify how to compute a suitable commutation relation. It characterizes the form of suitable commutation relations without providing a procedure for deriving them. It is in this sense that Method C is a search procedure. Once we have obtained a suitable commutation relation, it is clear from Dirac's principle that we can follow step C3 to derive a selection rule. However, we must first search for a suitable commutation relation. In contrast, the steps in Method G are decisive. Each step poses a question that has a definite answer. The Hamiltonian and perturbation operator have definite symmetries. Once we have determined these, we can implement the rest of the steps algorithmically. Hence,

group theory provides a superior structuring of the search space: it eliminates the need to search for a suitable commutation relation.

The group theoretic approach also provides enhanced tractability. Overall, the steps in Method G are easier to implement than the steps in Method C. Once we idealize the quantum mechanical system as possessing a well-defined symmetry group, the last five steps in Method G are straightforward computations. We simply identify the symmetries of the perturbation and initial wavefunction, compute their direct product, and apply Wigner's principle to determine a restriction on the allowed irreducible representations of the final state. Method C suffers from a deficiency in tractability due to step C2. It is difficult to determine a suitable commutation relation satisfying this constraint. Provided that we had a suitable commutation relation, it would be straightforward to apply step C3 and determine a selection rule. However, step C2 represents a tractability bottleneck: it provides little guidance on how to determine a suitable selection rule. In this case, tractability seems to mirror whether or not the method is a search procedure or decision procedure, but this need not always be the case. Sometimes search procedures are easier to implement than decision procedures, meaning that a search procedure can be more tractable than a decision procedure.

The commutator approach's deficiency in tractability points toward another organizational disadvantage: a lack of motivation. In many presentations of the commutator derivation of the selection rule for the orbital angular momentum, a suitable commutation relation is introduced with little or no exposition regarding why this commutation relation is suitable, or how one might originally discover that it is suitable. For instance, in Condon and Shortley's (1935, p. 60) and Griffiths' (2005, p. 373) presentations, no principled reasons are given for utilizing this commutation relation beyond the fact that it yields the correct result. This is not to say that a suitable commutation relation cannot be motivated. Heisenberg's presentation (1926, p. 369-371) motivates a similar commutation relation via an analogy with classical mechanics. Furthermore, as we saw at the end of Section 2.2.3, a motivational story can be told for this commutation rule. Nevertheless, Dirac's principle—as it stands—has difficulty motivating a suitable commutation relation. In contrast, Method G makes evident why we perform each step the way we do. The tools we use are motivated by Wigner's principle and the symmetry of the Hamiltonian. Getting clear on the significance of motivation and how alternative expressive means influence motivation is a difficult task. In the context of her own

case study, Rebecca Morris has developed a more thorough treatment than the cursory remarks I make here (unpublished).

Comparing the commutator and group theoretic approaches, their organizational differences are apparent. I have analyzed these differences in terms of organizational virtues, including enhanced modularization, effective structuring of the search space, tractability, and motivation. These virtues are organizational because they stem from how an expressive means structures a solution strategy for solving a problem. The commutator and group theoretic approaches differ in how they relate the operative physical concepts used in deriving selection rules. Group theory provides linguistic resources that enable these physical relations to be reorganized into a more effective solution strategy. An analogous structuring is not accessible from the commutator point of view, which lacks suitable expressive means to provide these organizational advantages. These organizational virtues manifest one sense in which the group theoretic approach provides a deeper understanding of selection rules.

4.3.2 Uniformity of Treatment and Unification

Distinct from modularization and tractability, we can also analyze the sense in which these approaches lead to a uniform treatment of selection rules. In this section, I characterize this notion of uniformity of treatment and compare it to Kitcher's account of unification (1989). I argue that the present case study motivates a distinction between uniformity of treatment and unification. Finally, I consider generality and its relationship to global fruitfulness. Each of these theoretical virtues represents an organizational advantage that I claim contributes to understanding.

Both the commutator and group theoretic approaches provide uniform treatments of selection rules in atomic spectroscopy. To provide a uniform treatment of a class of phenomena is to show how these phenomena can be understood via the same procedure. Provided we can derive suitable commutation relations, or alternatively characterize our system in terms of a suitable symmetry, either approach yields the relevant class of selection rules. This notion of uniform treatment is closely related to Kitcher's (1989) account of unification. According to Kitcher, a scientific theory unifies a class of phenomena when it derives those phenomena from the same underlying argument pattern. In Kitcher's terms, Methods C and G above are argument

patterns that draw upon the explanatory store of quantum mechanics and—in the case of Method G—group representation theory. On Kitcher's view, both approaches unify selection rules in atomic spectroscopy because they derive these selection rules according to the same argument pattern. Thus, Kitcher's notion of unification is analogous to the concept of uniform treatment that I adopt here.

The present case study suggests a sense of unification more restrictive than Kitcher's, one that is distinct from the notion of uniformity of treatment. An approach *uniformly treats* two explananda if it derives those explananda according to the same procedure. An approach *unifies* these two explananda if it shows how a single argument accounts for them both. The following example illustrates this distinction. The group theoretic approach provides a uniform treatment of both the selection rule on the total angular momentum and on the parity of electronic states (Laporte's rule). This means that the same argument pattern explains these rules, namely an instantiation of Method G. In one sense, this uniform treatment unifies these two rules by showing that they both follow from the same argument pattern—this is the sense of unification found in Kitcher. Yet, there is also a stronger sense in which group theory unifies the total angular momentum selection rule and Laporte's rule: it allows us to take these two separate instantiations of Method G and combine them into one derivation. Whereas the separate derivation of the total angular momentum selection rule utilizes the symmetry of the group SO(3), the separate derivation of Laporte's rule uses the symmetry of the inversion group. By taking the direct product of the special orthogonal group in three dimensions with the inversion group, we arrive at the mathematical structure known as the full orthogonal group (rotation-inversion group) in three dimensions: O(3). By utilizing the irreducible representations of O(3), we can derive the selection rules for the total angular momentum and Laporte's rule in a single unified derivation. This single derivation constitutes a unification of these two selection rules. If Wigner's remarks are to be heeded, the application of group theory was of great historical significance in explaining this rule. Wigner held the "explanation of Laporte's rule (the concept of parity)" to be one of the most significant contributions of the application of group theory to quantum mechanics (1959, v).

Intriguingly, this unification extends even further. From the rows of the irreducible representations of the full orthogonal group we arrive at the selection rules for the z-component of the angular momentum. Furthermore, we can even build in the permutation symmetries of the

electrons into the group of the Schrödinger equation, allowing us to derive—in one unified derivation—the selection rule on the spin quantum number. Thus, through a more expansive view of the group of the Schrödinger equation, the group theoretic approach affects a unification of selection rules in atomic spectroscopy. This provides a meaningful sense of unification that can be distinguished from uniformity of treatment. On this dimension, the group theoretic approach is superior to the commutator approach. The commutator approach lacks suitable expressive means to effect an analogous unification. In principle, we could concatenate the commutation relations for the various selection rules into a longer commutation relation. If we had this overarching commutation relation, we could derive the selection rules in one argument structure. However, I contend that the algebra involved in this derivation would be intractable: we would not be able to derive the selection rules. Furthermore, this purported "unification" would be artificial in the sense that no clear physical meaning would attach to the concatenated commutation relation. This is in contrast to the mathematical structure of the full orthogonal group, which has a clear interpretation as an underlying symmetry of atomic systems.

Uniform treatments and unifications enhance understanding by casting multiple phenomena under a single viewpoint. Expressive means that recast phenomena uniformly enable phenomena to be understood uniformly. Uniformity is closely related to the notion of generality, which considers the scope of an intellectual framework. By definition, uniform treatments are general: they treat multiple phenomena uniformly. However, uniform treatments vary in generality. Some uniform treatments handle proper subsets of phenomena that are covered by more general uniform treatments. Within the context of atomic spectroscopy, both Methods C and G are equally general: they cover the same range of phenomena. Once we step outside the bounds of the present case study and consider selection rule phenomena in molecular and nuclear spectroscopies, the group theoretic approach gains the upper hand. The group theoretic method supplies a uniform treatment in terms of molecular and nuclear symmetries. In contrast, it is not clear that appropriate commutation relations can be found in order for the commutator approach to apply in these areas. If this is correct, then the group theoretic account provides a more general treatment of selection rule phenomena. This would support the claim—which again goes beyond this case study—that the group theoretic account is globally more fruitful than the commutator account.

5.0 CONCLUSIONS

In Chapter 2, I presented two different methods for deriving selection rules: a commutator approach and a group theoretic approach. These methods are compatible with each other, so we can accept them both. Rather than focus on the fact *that* these approaches theoretically justify selection rules, I have examined *how* they justify selection rules. I have used the intellectual differences between the commutator and group theoretic approaches to motivate my positive accounts of scientific explanation and understanding. To conclude, I summarize how these accounts provide a philosophical picture of my case study. I clarify what I take to be the organizational nature of scientific understanding.

We began in Chapter 1 with the presumption that group theory provides a deeper understanding of selection rules than the commutator approach. My task has been to explicate this pre-theoretical notion of scientific understanding, providing a partial account of what it consists in. Before developing this account, I considered the nature of explanation in science. It is natural to require that understanding arises from explanatory arguments. Motivated by this relationship between explanation and understanding, Chapter 3 considered whether commutator and group theoretic derivations of selection rules should be viewed as explanatory or non-explanatory. According to a relevance account of explanation, these derivations are explanatory insofar as they draw upon physical and mathematical features that are relevant to selection rule phenomena, while eliminating irrelevant features. Since the commutator and group theoretic derivations draw upon different physical features, a strong interpretation of relevance accounts entails that only one approach can be explanatory. However, the underlying assumption of relevance accounts—namely that relevance grounds explanation—fails to describe scientific practice. As I argued in Section 3.3, there are cases where we determine whether a feature is relevant or irrelevant by providing explanations. Batterman and Rice's minimal model account of explanation raises a similar problem for relevance accounts.

In Section 3.5, I argued that the explanatory constraints supported by Batterman and Rice are likewise too strong. Scientists do not always provide explicit stability analyses of explanations of universal behavior. In the case of selection rules for atomic spectra, most—if not all—derivations shirk a rigorous stability analysis. Instead, these derivations implicitly assume that, for systems of low atomic number, perturbations disrupting the central field approximation are negligible. Hence, I recommended interpreting Batterman and Rice's constraints as guidelines for good explanatory practice, rather than strict requirements for explanations. I believe these constraints should be taken as dealing with the quality of an explanation as opposed to explanation *simpliciter*. Accordingly, the derivations of selection rules support a minimal account of mathematical explanations wherein arguments are explanatory if they recover the phenomenon of interest according to accepted physical and mathematical constraints. My account treats mathematical explanation as inherently justificatory: an explanation justifies a phenomenon. Although we could prescribe additional features as constraints on explanation, I believe it is more felicitous to view these features as constraints on understanding. On my account, the derivations of selection rules in Chapter 2 are explanatory, and we can naturally move to an analysis of how they differ in the understanding they provide.

My account of understanding relies on super-empirical virtues, such as modularization, tractability, uniformity of treatment, and unification. Unlike in the context of traditional underdetermination debates, it is irrelevant whether or not these virtues are epistemic indicators of truth-likeliness. In the context of compatible explanations, relative truth-likeliness is not in question. Rather than pertaining to truth or a metaphysical interpretation of our physical theories, these virtues are organizational in nature. They deal with how an expressive means structures solution procedures to problems. Alternative expressive means enable us to focus on different physical features, leading to differences in expressive power. Having the conceptual resources to analyze certain physical features—such as symmetry types—greatly facilitates problem-solving. My account of understanding provides a method for distinguishing the intellectual content of the commutator and group theoretic approaches. I argued that group theory provides a superior structuring of the search space when deriving selection rules. Group theory provides an intellectual advantage by re-expressing and reorganizing the mathematical and physical constraints that undergird these derivations. To effect this organization, we require sufficient expressive power to discuss the symmetry types of atomic systems. This, in turn, requires

sufficient expressive means. Through the concept of irreducible representations, group representation theory provides suitable expressive means to accomplish this task. I believe that my theoretical account of understanding captures salient features of our pre-theoretical notion of scientific understanding.

In this thesis, I have analyzed one example where group theory provides a superior understanding of atomic spectra. Although this necessarily limits the scope of my conclusions, I am optimistic that features of this case study will generalize to other cases, both within and outside atomic spectroscopy. Within the theory of atomic spectra, my methodology applies to the Stark and Zeeman effects, the Wigner-Eckart theorem, the coupling of angular momenta, and perturbation theory, where both non-group theoretic and group theoretic approaches have been developed. I plan to analyze these phenomena to strengthen my account of how group theory deepens scientific understanding of atomic spectra. Outside atomic spectra, non-group theoretic alternatives are not so readily found. Although lacking a benchmark for comparison complicates philosophical analysis, I am hopeful that lessons learned from atomic spectra will facilitate a more thorough analysis of group theory's role in molecular and nuclear spectroscopies and particle physics. This case study supports a crucial philosophical moral relevant in these other contexts: we cannot account for group theory as an intellectual advance merely through the explanations it provides. Recognizing that group theory provides explanations is only the start of a philosophical analysis. The real work consists in explicating *ways in which* group theory furnishes explanations. Lessons learned from group theoretic cases may further generalize to other areas of applied mathematics. I hope to have motivated the fruitfulness of taking a closer look at how mathematics contributes to scientific understanding.

BIBLIOGRAPHY

Batterman, R. W. (2002). *The devil in the details : asymptotic reasoning in explanation, reduction, and emergence* (Oxford studies in philosophy of science). New York: Oxford University Press.

Batterman, R. W. (2010). On the explanatory role of mathematics in empirical science. *British Journal for the Philosophy of Science, 61*, 1-25.

Batterman, R. W., & Rice, C. C. (unpublished). Minimal model explanations.

Born, M., Heisenberg, W., & Jordan, P. (1926). Zur Quantenmechanik II (B. L. v. d. Waerden, Trans.). In *Sources of Quantum Mechanics* (pp. 557-615, Classics of Science). New York: Dover.

Born, M., & Jordan, P. (1930). *Elementare Quantenmechanik*. Berlin: Springer.

Borrelli, A. (2009). The emergence of selection rules and their encounter with group theory, 1913-1927. *Studies in History and Philosophy of Modern Physics, 40*, 327-337.

Condon, E. U., & Shortley, G. (1935). *The theory of atomic spectra*. Cambridge Eng.: The University Press.

Cornwell, J. F. (1984). *Group theory in physics* (Techniques of physics, Vol. 7). Orlando: Academic Press.

Dirac, P. A. M. (1930). *The principles of quantum mechanics* (The international series of monographs on physics). Oxford: The Clarendon Press. (2nd ed.: 1935; 3rd ed.: 1947; 4th ed.: 1958).

Engel, T., & Hehre, W. J. (2010). *Quantum chemistry & spectroscopy* (2nd ed.). New York: Prentice Hall.

Flurry, R. L. (1980). *Symmetry groups: theory and chemical applications*. Englewood Cliffs, N.J.: Prentice-Hall.

Griffiths, D. J. (2005). *Introduction to quantum mechanics* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Güttinger, P., & Pauli, W. (1931). Zur Hyperfeinstruktur von Li+. Teil II, Mathematischer Anhang. *Zeitschrift für Physik, 87*, 754-765.

Hempel, C. G. (1965). Aspects of scientific explanation. In *Aspects of scientific explanation and other essays in the philosophy of science*. (pp. 331-496). New York: Free Press.

Herzberg, G. (1944). *Atomic spectra and atomic structure*. (*Trans.* J. W. T. Spinks). New York: Dover publications.

Judd, B. R. (1963). *Operator techniques in atomic spectroscopy* (McGraw-Hill advanced physics monograph series). New York: McGraw-Hill.

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher, & W. Salmon (Eds.), *Scientific Explanation* (Vol. 13, pp. 410-505, Minnesota studies in the philosophy of science). Minneapolis: University of Minnesota.

Manders, K. (unpublished). Expressive means and mathematical understanding.

Morris, R. (unpublished).

Petrashen, M. I., & Trifonov, E. D. (1969). *Applications of group theory in quantum mechanics*. Cambridge, Mass.,: M.I.T. Press.

Petrashen, M. I., & Trifonov, E. D. (2009). *Applications of group theory in quantum mechanics* (Dover ed.). Mineola, N.Y.: Dover Publications.

Pincock, C. (2011). On Batterman's 'On the explanatory role of mathematics in empirical science'. [discussion note]. *British Journal for the Philosophy of Science, 62*, 211-217.

Tinkham, M. (1964). *Group theory and quantum mechanics* (International series in pure and applied physics). New York: McGraw-Hill.

Tinkham, M. (2003). *Group theory and quantum mechanics*. (Dover ed.). Mineola, N.Y.: Dover Publications.

Waerden, B. L. v. d. (1932). *Die gruppentheoretische Methode in der Quantenmechanik* (Die Grundlehren der mathematischen Wissenschaften, Vol. Bd XXXVI). Berlin: J. Springer.

Weyl, H. (1927). Quantenmechanik und Gruppentheorie. *Z. Physik, 46*, 1-46.

Weyl, H. (1928). *Gruppentheorie und Quantenmechanik*. Leipzig: S. Hirzel.

Weyl, H. (1931). *The theory of groups and quantum mechanics*. (*Trans.* H. P. Robertson). London: Methuen & co. ltd.

Wigner, E. (1927). Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen. *Z. Physik, 67*, 624-652.

Wigner, E., & Neumann, J. v. (1928a, b, c). Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons. *Z. Physik, 47, 49, 51*, 203-220, 273-294, 844-858.

Wigner, E. P. (1931). *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Die Wissenschaft Sammlung von Einzeldarstellungen, Vol. Bd 85). Braunschweig: F. Vieweg & Sohn Akt.-Ges.

Wigner, E. P. (1959). *Group theory and its application to the quantum mechanics of atomic spectra* (Expanded and improved ed., Pure and applied physics, Vol. 5). New York: Academic Press.