SHATTERED SYMMETRY

GROUP THEORY FROM THE EIGHTFOLD WAY TO THE PERIODIC TABLE

Pieter Thyssen

and

Arnout Ceulemans

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

Library of Congress Cataloging-in-Publication Data
Names: Thyssen, Pieter. | Ceulemans, Arnout.
Title: Shattered symmetry : group theory from the eightfold way to the periodic table / by Pieter Thyssen and Arnout Ceulemans.
Description: Oxford : Oxford University Press, 2017. | Includes bibliographical references and index.
Identifiers: LCCN 2016017431 | ISBN 9780190611392
Subjects: LCSH: Group theory. | Symmetry (Physics) | Lie algebras. | Logic, Symbolic and mathematical. | Periodic table of the elements.
Classification: LCC QC20.7.G76 T49 2016 | DDC 512/.2–dc23 LC record available at https://lccn.loc.gov/2016017431

9 8 7 6 5 4 3 2 1
Printed by Sheridan Books, Inc., United States of America

Contents

Part One Space Symmetries

Part Two Dynamic Symmetries

Part Three Spectrum-Generating Symmetries

Figures

Tables

Preface

Confronted with the bewildering complexity of the world we inhabit, scientists of all ages have attempted to bring order and simplicity to this apparent chaos. Baffled by the potpourri of chemical elements, the zoo of elementary particles, and the plethora of spectral lines in hydrogen's emission spectrum, scientists have looked high and low for some underlying unity. To this end, the power of symmetry arguments cannot be overemphasized.

Hidden behind the variegated phenomena of every day lies a crystalline perfection that has shattered to give birth to the real world. As the Platonic world of grand mathematical symmetries crumbled over time, our asymmetric reality materialized. But with this spontaneous breaking of symmetry, patterns started to emerge as well. The act of bifurcation gave rise to variety and identity. "*C'est la dissymétrie qui crée le phénomène*," dixit Pierre Curie.[1]

This textbook deals not only with the obvious discrete symmetries of crystals and molecules, but is also concerned with the deeper lying continuous symmetries that are at the basis of the dynamics of physical systems. The description of these symmetries is made possible by the mathematical theory of groups. *Shattered Symmetry* consists of three parts that build up the knowledge of groups gradually.

The first part describes the legacies of Évariste Galois and Sophus Lie. Galois introduced the concept of a discrete group and Lie defined the algebras that describe continuous groups. Our discovery of these concepts is based on the perfect symmetries of the circle and the sphere, which visualize Lie group symmetry directly. The part closes with a Scholium, which introduces the representation of groups by means of Cartan–Weyl diagrams. These are an important graphical tool to classify and label states of symmetric systems used throughout the book.

The second part deals with the two most important central field problems in physics: the spherical oscillator and the Coulomb hole. Both potentials are characterized by high dynamic symmetries that go beyond spherical geometry and render the corresponding dynamical problems exactly solvable. During this process the associated algebras generate the quantum numbers that characterize all the states of the physical system. The oscillator problem embodies unitary symmetry and is at the basis of the Eightfold Way of Murray Gell-Mann, which brought order to the particle zoo in elementary particle physics. The particle in a Coulomb hole is described by the four-dimensional rotation group, which corresponds to the dynamical symmetry of the hydrogen atom and thus forms the cornerstone of atomic

---

[1] P. Curie. "Sur la Symétrie dans les Phénoménes Physiques, Symétrie d'un Champ Électrique et d'un Champ Magnétique." *Journal de Physique Théorique et Appliquée* 3.1 (1894), pp. 393–415. Quote on p. 400.

physics. Both these groups are of paramount importance because they encode the elementary quantum states of matter. Again, a Scholium at the end of this group of chapters provides a convenient resting place to look deeper into the connection between unitary groups and rotations.

In the third part, we leave the shores of compact Lie algebras and set course to the continent of the noncompact Lie algebras, which unite all spectral levels of a physical system into one manifold and lead to the construction of the comprehensive SO(4,2) symmetry of hydrogen, which collects the quantum states into one symmetry space. In the final chapters of this book, this group theoretical construction is confronted with the periodic system of Mendeleev, which gives all the chemical elements a place. At this point, the perfect symmetry of the hydrogen atom is transformed, broken and tilted, to give rise to the intriguing Aufbau of the periodic table, with its strange period doubling and Madelung regularity.

The presentation is inspired by a naive "chemical" approach. This means the various Lie algebras are viewed as distinct "molecules", with their own stories and individuality. Accordingly, the logic of the presentation follows a line of parentage, taking us from one group to the next. In Part I, we start with the symmetry of the circle, described by the simple SO(2) group, and then raise it to the three dimensions of spherical symmetry described by the SO(3) group. Both these groups have the advantage that they can be visualized easily because they relate to simple geometric objects with a perfect symmetry. In Part II, this easy link with geometry is lost as we extend our chain of groups from SO(3) to the unitary group SU(3), which in many respects can be treated as SO(3)'s bigger brother. The treatment of SU(3) in Chapter 7 brings us to the first summit of our journey, from which a magnificent panorama on elementary particles is deployed. We then descend, simply by taking out the third dimension and reach an SU(2) plateau. From there, the road rises up again to reach the SO(4) dynamic symmetry group of hydrogen, which is nothing else than the product of SU(2) × SU(2). At the horizon, we finally see a glimpse of the Periodic Table, but we are not there yet. In Part III, we first have to open up the Lie groups to accommodate the infinite manifold of the entire spectrum of a physical system. We first deal with the SO(2,1) group, which is the noncompact version of SO(3). Subsequently, by combining this SO(2,1) with the SO(4) group of hydrogen, we have all the ingredients to construct SO(4,2). This group defines the board on which we are going to play chess that finally reveals the structure of Mendeleev's Table.

Our greatest concern in this book has been to maintain a high degree of accessibility throughout. Theoretical concepts are explained in a straightforward, intuitive manner; all calculations are worked out fully in appendices, and mathematical concepts (e.g., matrix and vector algebra) are reviewed briefly; most chapters commence with a historical introduction.

Our greatest satisfaction would be to succeed in presenting Lie algebras in a self-contained way, which requires only very little mathematical background, without betraying, however, the beautiful and deep mathematics that is their basis.

Hopefully, readers are inspired from time to time by Tenniel's vintage illustrations from Alice in Wonderland,[2] which are scattered throughout the book as metaphors of group theoretical concepts.

---

[2] For Lewis Carroll's major tales, *Alice's Adventures in Wonderland* and *Through the Looking Glass*, with the original illustrations by John Tenniel, see L. Carroll. *The Annotated Alice: The Definitive*

During the course of the five years it took to create this book, we had several opportunities to discuss various aspects with colleagues and friends. We gratefully acknowledge contributions from Christiane Quesne, Eric Scerri, Eugene Schwartz, Olimpia Lombardi, Rutwig Campoamor-Stursberg, and many more. We are grateful to many proofreading volunteers, especially to Tohru Sato for a thorough check of Parts I and II. Yutaka Kitagawara shared with us his ideas on nonlinear Lie algebras. We are indebted to Loïc Casson and "la Société Académique de l'Oise, Beauvais, France" for providing valuable information on the life and work of Charles Janet. We thank Dalibor Hršak for the figure of the Möbius hexaphyrin, and Maria Daikidou for her translation of a phrase in Plato's *Timaeus*. Daniela Di Lena kindly granted permission to use her artwork of chess pieces on the front cover.

Edition. With Introduction and Notes by Martin Gardner, and Illustrations by John Tenniel. New York: W.W. Norton, 1999.

PART ONE

SPACE SYMMETRIES

1 A primer on symmetry

What was there in the beginning? A physical law, mathematics, symmetry? In the beginning was symmetry!
–Werner Heisenberg[1]

1.1 THE TRAGIC LIFE OF ÉVARISTE GALOIS

Bang! Early in the morning on Wednesday, May 30, 1832, the loud crack of a single gunshot filled the air around the Glacière pond in Gentilly, near Paris. A local peasant,[2] who was bringing vegetables to market that day, rushed toward the sound and found a young man lying on the ground, fatally shot in the abdomen during a duel. The young man was Évariste Galois (1811–1832), a twenty-year-old mathematical genius and well-known revolutionary in Paris (Figure 1.1).[3]

---

[1] W. Heisenberg. "Physics and Beyond: Encounters and Conversations." In: *World Perspectives*, vol. 42. Ed. R. N. Anshen. New York: Harper & Row, 1971, p. 132.

[2] He might have been a former army officer; no one really knows.

[3] *Biographical notes*: Several French biographies of Galois have appeared over the years. The first biography was written by P. Dupuy. "La Vie d'Évariste Galois." In *Annales Scientifiques de l'École Normale Supérieure*, 3rd ser.13 Paris: Gauthier-Villars, 1896, pp. 197–266 (reprinted in P. Dupuy. *La Vie d'Évariste Galois*. Paris: Éditions Jacques Gabay, 1992). Other French works include A. Dalmas. *Évariste Galois, Révolutionnaire et Géomètre*. Paris: Fasquelle, 1956; A. Astruc. *Évariste Galois*. Paris: Flammarion, 1994; N. Verdier. *Évariste Galois, le Mathématicien Maudit*. (Collection: *Les Génies de la Science*. No 14.) Paris: Pour la Science, 2003; and J.-P. Auffray. *Évariste 1811–1832, Le Roman d'une Vie*. Lyon: Aléas, 2004. Some of the best-documented English biographies include G. Sarton. "Evariste Galois." *Scientific Monthly* 13.4 (1921), pp. 363–375 (reprinted in G. Sarton. "Evariste

FIGURE 1.1 Portrait of Évariste Galois (1811–1832), drawn from memory by his brother Alfred in 1848.

Although his intestines and stomach had been perforated by the bullet, Galois was still alive and was brought to the Côchin hospital around half past nine. His younger brother Alfred hurried to visit him and started crying. "Don't cry Alfred," said Galois, trying to console his brother with his last words. "I need all my courage to die at twenty." That evening, peritonitis set in and, at ten o'clock the next morning, Galois passed away.[4]

Galois." *Osiris* 3 (1937), pp. 241–259), T. Rothman. "Genius and Biographers: The Fictionalization of Évariste Galois." *American Mathematical Monthly* 89.2 (1982), pp. 84–106 and L. Toti Rigatelli. *Evariste Galois: 1811–1832.* (*Vita Mathematica*, vol. 11.) Ed. E. A. Fellmann. Basel: Birkhäuser Verlag, 1996. Shorter accounts of Galois' life and work include those found in A. Chevalier. "Nécrologie Evariste Galois." *Revue Encyclopédique* 55 (1832), pp. 744–754 and T. Rothman. "The Short Life of Évariste Galois." *Scientific American* 246.4 (1982), p. 136–149. Galois' tragic life has also been the focus of interest in popular scientific literature. See especially M. Livio. *The Equation That Couldn't Be Solved: How Mathematical Genius Discovered the Language of Symmetry*. New York: Simon & Schuster Paperbacks, 2006, pp. 112–157; but also see E. T. Bell. *Men of Mathematics*. Melbourne: Penguin Books, 1953, pp. 398–415; I. James. *Remarkable Mathematicians: From Euler to von Neumann*. Cambridge: Cambridge University Press, 2002, pp. 134–141; I. Stewart. *Why Beauty Is Truth: A History of Symmetry*. New York: Basic Books, 2007, pp. 97–123; and M. Ronan. *Symmetry and the Monster: One of the Greatest Quests of Mathematics*. Oxford: Oxford University Press, 2006, pp. 11–26.

[4] The circumstances of this tragic affair defy rational explanation and have been veiled in mystery to this day. Contradictory claims abound and, as Tony Rothman piercingly noted, most accounts of Galois' untimely death are "baroque, if not byzantine, inventions." See T. Rothman. *Science a la mode. Physical Fashions and Fictions*. Princeton: Princeton University Press, 1989, p. 186. Various biographers (Galois' brother included) are convinced Galois had been murdered by political enemies. Galois, however, apologized in his letter "to all republicans" for dying "the victim of an infamous coquette"—alluding to his failed love affair with Stéphanie Potterin du Motel in spring 1832. See T. Rothman. "The Short Life of Évariste Galois." 1982, p. 149. Infeld's hypothesis that Stéphanie was, in fact, a prostitute who had been hired by the police seems a little far-fetched. See L. Infeld. *Whom the Gods Love: The Story of Evariste Galois*. New York: Whittlesey House, 1948. In any case, Galois mentioned *two* contestants in the duel and emphasized they were "two patriots" who had acted in "good faith." In line with this, Gabriel Demante (a cousin of Galois) noted the duel had been provoked by Stéphanie's uncle and fiancé. Yet another theory was recently put forth by the

1.1.1 Entrance exams

A few years earlier, while still a student at the Lycée Louis-le-Grand, Galois had become spellbound by the theory of equations. Having devoured Adrien-Marie Legendre's (1752–1833) *Eléments de Géométrie* in just two days, Galois began to read one professional mathematics paper after another from the pens of Joseph-Louis Lagrange (1736–1813) and Niels Henrik Abel (1802–1829).

Full of confidence, Galois attempted to enter the famous École Polytechnique, one of the most prestigious institutes in Paris. Barely sixteen years old, completely unprepared and too lazy to document his own reasoning, it was not surprising Galois failed the entrance exam. When he tried to gain entrance a second time in August 1829, his father had just committed suicide by suffocating himself. Galois, who was still mourning his father's death, was unable to think clearly, and lost his temper when one of the examiners started asking questions he deemed childishly simple. The impatient student threw a blackboard eraser into the examiner's face, and lost his final chance of being accepted at the École.

1.1.2 Publish or perish

But, Galois did not give up easily. Building on the work of Lagrange and Abel, he had recently found a way to resolve a great mathematical riddle, one that had tormented the mathematical community for centuries, and he was determined to disseminate his ideas. Toward the end of May 1829, Galois submitted a paper to the French Academy of Sciences that outlined his great breakthrough. Augustin-Louis Cauchy (1789–1857), France's leading mathematician at the time, was appointed as a referee, but kept the manuscript for more than six months without ever presenting Galois' research at a meeting of the Academy.[5]

Cauchy must have been impressed by Galois' work though, as he advised the young man to expand on his ideas and to resubmit an enlarged manuscript for the Grand Prix in Mathematics. A few days before the deadline, in February 1830, Galois sent his revised memoir to the Academy. Joseph Fourier (1768–1830), the perpetual secretary of the Academy, took the manuscript home, but died soon after, on May 16, 1830. Galois' memoir got lost and was not recovered in time to be considered for the Grand Prix.

Furious, Galois resubmitted his manuscript for a third time on January 17, 1831. By then, Cauchy had fled Paris because of the political turmoil in the capital and Fourier was dead. So, Galois' paper fell into the hands of Siméon Denis Poisson (1781–1840) and Sylvestre François Lacroix (1765–1843). Two months passed without any response and Galois was edging slowly toward a nervous breakdown. Finally, on July 4, 1831, Poisson responded, but it was a crushing blow. Poisson judged Galois' memoir to be utterly "incomprehensible" and unfit for publication.

Italian mathematician and historian Laura Toti Rigatelli. Rigatelli believes the disillusioned Galois sacrificed himself by staging the entire duel in an attempt to stir up rebellion. Given his two failed attempts to enter the École Polytechnique, the triple rejection of his first memoir, and his broken heart, Galois did not care to live any longer. Wherever the truth may lie, the details surrounding Galois' duel are still shrouded in mist, and we leave it to you to make up your own mind.

[5] On January 18, 1830, Cauchy wrote an apologetic letter to the Academy. He was expected to read Galois' paper and one of his own that day, but was feeling slightly unwell, and decided to stay home. At the next meeting, however, on January 25, Cauchy presented only his own paper.

1.1.3 Galois' mathematical testament

On May 29, 1832, the evening of the duel, Galois realized he was probably going to die the next morning, and this was his very last chance to try to explain his great breakthrough. So he stayed up all night, writing farewell letters to his friends, and composing what would become his mathematical testament. Galois' *Lettre Testamentaire* comprised seven hastily written pages summarizing his latest achievements. The letter was addressed to his close friend, Auguste Chevalier (1809–1868), and Galois urged him to take care of publishing his manuscripts. Because Poisson had failed to understand Galois' first memoir, Galois spent the rest of the night adding last-minute corrections to his first memoir (Figure 1.2). Time was running short, however, and at one point Galois broke off to scribble in the margin, "Je n'ai pas le tem[p]s" ("I have no time"), before he moved on to the next paragraph. The sun's first rays had already pierced the sky when Galois went to meet his destiny, leaving a small pile of papers in the middle of his desk.

Contained inside these documents was a new language, a language that would enable the scientific community finally to understand one of the most fundamental concepts of Nature—namely, *symmetry*. And that language, a real calculus of symmetry, was called *group theory*.

The aim of this book is to learn how to *speak* group theory. But, before we can do any calculus on symmetry, we must first acquaint ourselves with the concept of symmetry itself.

FIGURE 1.2 Galois' *Premier Mémoire* contained the seeds of modern group theory. On the eve before his duel, Galois revised his manuscript and added some last-minute marginal additions. One of them was the dramatic sentence "Je n'ai pas le tem[p]s" ("I have no time"), as can be seen on the left-hand side of the manuscript.

1.2 THE CONCEPT OF SYMMETRY

Symmetry lies at the heart of the world. From the six-sided snowflakes in winter to the floral patterns in summer, symmetry is all around us. Its myriad manifestations are overwhelming; there is *rotational symmetry* in flowers, *helical symmetry* in seashells, *spiral symmetry* in the Pinwheel galaxy, *cubic symmetry* in salt crystals, and *bilateral symmetry* (i.e., symmetry of left and right) in the human body, as exemplified in Da Vinci's *Vitruvian Man* (Figure 1.3).

In addition, symmetry also inspires; it stirs our imagination and awakens our creativity. Think of Escher's masterpieces or the stories of Carroll, the fugues of Bach

FIGURE 1.3 Leonardo Da Vinci's *Vitruvian Man*, illustrating the bilateral symmetry of the human body (*Gallerie dell Academia*, Venice, 1492).

FIGURE 1.4 Symmetry provides a key to Nature's secrets, which opens the door to a deeper understanding of the natural phenomena that surround us.

or Balanchine's ballet choreographies. And yet, perhaps most important, symmetry intrigues us; it stimulates our thinking and drives us onward in our quest to understand.

Not surprisingly, then, symmetry has become a central principle in science, underlying many of the revolutionary concepts of modern physics and chemistry. As Steven Weinberg (Nobel laureate in Physics) recently exclaimed, symmetry provides a key to Nature's secrets (Figure 1.4).[6] It unlocks the door to a profound understanding of the physical world and all the wonders it entails.

1.2.1 Symmetry defined

Before we continue, let us attempt to *define* symmetry. Humankind has been familiar with the concept of symmetry for thousands of years, but although we all share an intuitive feeling for what symmetry is all about, most of us would have a hard time coming up with a precise, succinct definition.[7] We have no problem differentiating symmetric objects from asymmetric ones, but are mute when asked to explain the distinction. The same worriment befell early Christian theologian and philosopher St. Augustine (354–430) when he was challenged to define the concept of time. "If no one asks me, I know," he wrote in his *Confessions*.[8] "But if I wish to explain it to one that asketh, I know not." Not surprisingly, then, the modern definition of symmetry took hundreds of years to develop.

The ancient Greek definition of symmetry

The first definition of symmetry arose in Ancient Greece around the fourth century BC. Despite its age and its overthrow by more modern definitions, this primitive

---

[6] S. Weinberg. "Symmetry: A 'Key to Nature's Secrets.' " In: *The New York Review of Books*, vol. 58.16. Ed. R. B. Silver. New York: Rea S. Hederman, 2011, p. 69. See also S. Weinberg. "Varieties of Symmetry." *Symmetry: Culture and Science* 23.1 (2012), pp. 5–16.

[7] Just try it for yourself and you'll notice how challenging it actually is.

[8] St. (Bishop of Hippo) Augustinus, Book 11, Chapter 14, circa the year A.D. 397.

and rather subjective definition still lies closest to our heart. The Greek philosophers associated symmetry with *harmony* and *balance*, with *regularity* and *order*; they introduced the word συμμετρια to convey that something was *well-proportioned* or of the *right balance* (literally, "with measure," from συν + μετρον). Symmetry was also considered a sign of *beauty* and *perfection*; symmetric objects were said to be aesthetically appealing.

In Aristotle's (384–322 BC) words, for example, "the chief forms of beauty are order (τάξις) and symmetry (συμμετρια) and definiteness (ω̇ρισμενον), which the mathematical sciences demonstrate in a special degree."[9] Another example comes from Plato's (428–348 BC) *Timaeus*, which recounted how the Universe had been created out of chaos by a Demiurge using the four elements fire, air, earth, and water as basic building blocks. These elements were associated with four of the five *Platonic solids*—the tetrahedron, octahedron, cube, and icosahedron—which were considered to be symmetric in the sense of being well-proportioned, regular, and aesthetic (Figure 1.5).

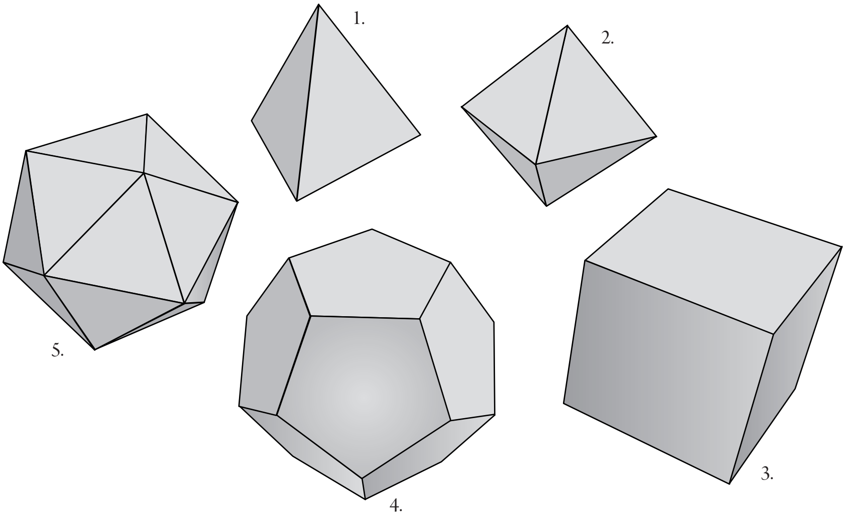


FIGURE 1.5 The mathematical beauty and symmetry of the five Platonic solids: 1, tetrahedron; 2, octahedron; 3, cube (or hexahedron); 4, dodecahedron; and 5, icosahedron. The element fire was identified with the pointy tetrahedron; air with the smooth octahedron; earth with the bulky, weighty cube; and water with the fluid and nearly spherical icosahedron. Aristotle later added the dodecahedron to this list, and postulated that it represented the *aether*—a quintessential substance that made up the celestial heavens.[10]

---

[9] Aristotle. *Aristotle's Metaphysics, a Revised Text with Introduction and Commentary by W. D. Ross*. Oxford: Clarendon Press, 1924. Book 13, 1078.a3. Hermann Weyll (1885–1955) echoed this in his little book on symmetry. "Symmetry," he said, "as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty and perfection." See H. Weyl. *Symmetry*. Princeton: Princeton University Press, 1952, p. 5.

[10] In his *Timaeus*, Plato devoted only one sentence to the dodecahedron, the meaning of which is ambiguous. In one version it reads: "There remained one construction, the fifth; and the God used it for the whole, making a pattern of animal figures thereon." See F. Mc. D. Cornford. *Plato's Cosmology*. London: Routledge, 1937, reprinted 2014, p. 218. A dodecahedron has twelve faces and Plato most probably believed the Demiurge had used this solid for the creation of the entire

The modern-day definition of symmetry

The question remains, however, *how* all this can be translated into mathematically precise terms. Harmony, balance, beauty, and perfection are all vague concepts, and it is far from clear how they relate to physics. Indeed, according to Hermann Weyl (1885–1955), “The Greeks never used the word ‘symmetric’ in our modern sense.”[11] For centuries, scientists (and artists alike) continued to use the term *symmetry* in the previously mentioned hand-waving sense, while its modern scientific connotation lay dormant.[12]

Only during the late eighteenth century did scientists come up with the first modern definition of symmetry. The idea was really very simple, and can be illustrated with the following excerpt from Lewis Carroll’s *The Hunting of the Snark*:[13]

You boil it in sawdust: you salt it in glue:
You condense it with locusts and tape:
Still keeping one principal object in view—
To preserve its symmetrical shape.

Despite its literary nonsense, these lines nonetheless contain the essence of symmetry. Carroll is writing about a mysterious object that is boiled, salted, and condensed. And yet, although the object is acted on in so many different ways, it somehow retains its original shape. This “immunity to a possible change” brings us to the scientific definition of symmetry.[14]

Definition 1.1 (Symmetry): A (mathematical) object is said to be *symmetric*, or to possess *a symmetry*, when there is a *transformation* that leaves the appearance of the object unchanged. ■

There are two crucial ingredients to this definition[15]:

1. *Possibility of change.* It should be possible to transform the object.
2. *Immunity to change.* When transformed, some of the object’s features should remain unchanged.

Universe along with the zodiac and its twelve constellations. An alternative version translates as: “this the Demiurge used in the delineation of the Universe.” The Greek word for *delineation* is διαζωγραφῶν, which translates as “painting in semblance of life,” but it may also refer to the depiction of living beings as the animals of the zodiac. See R. F. Kotrc. “The Dodecahedron in Plato’s Timaeus.” *Rheinisches Museum für Philologie* 124 (1981), pp. 212–222.

[11] Weyl, *Symmetry*, p. 75.

[12] The great Renaissance artist Leonardo da Vinci (1452–1519), for instance, emphasized the symmetric proportions of the human body (known as the *Canon of Proportions*) in his *Vitruvian Man* around 1490 (Figure 1.3).

[13] L. Carroll. “The Beaver’s Lesson.” In: *The Hunting of the Snark: An Agony in Eight Fits*. London: Macmillan, 1876, p. 56.

[14] J. Rosen. *Symmetry in Science: An Introduction to the General Theory*. New York: Springer-Verlag, 1995, p. 2; M. Livio. *The Equation That Couldn’t Be Solved: How Mathematical Genius Discovered the Language of Symmetry*. New York: Simon & Schuster, 2006, p. 4.

[15] Rosen, *Symmetry in Science*, p. 4; J. Rosen. *Symmetry Rules: How Science and Nature Are Founded on Symmetry*. The Frontiers Collection. Berlin: Springer, 2008, p. 4

Notice that we are no longer defining symmetry in *passive*, *static* terms (such as *perfection* and *balance*), but in *active*, *dynamic* terms. Symmetry is no longer a *thing* to be perceived in the objects around us; it has become a *process*, a special kind of action—a way to *transform* the object such that its structure is preserved.[16]

When seen from this operational perspective, symmetry is very much like a magician's trick. A great illusionist shows you a particular object and asks you to close your eyes for a moment. At that point, he does something to the object; he transforms it in a certain way. If, when you open your eyes, the object looks exactly as it did before and you are absolutely unable to tell whether the magician has fiddled around with it, then that transformation was a symmetry of the object.[17]

Summarizing, we could say that to describe the symmetry of an object means to elucidate the nature of identity in the context of change. Symmetry is all about this never-ending quest for permanence in a world of constant flux.

1.2.2 The symmetries of a triangle

Consider, by way of introductory example, the equilateral triangle in Figure 1.6.[18] Which transformations of the triangle leave its appearance (in terms of size, shape, and orientation) apparently unchanged? Imagine holding a cardboard triangle and placing it in front of you on the table while drawing an outline around it. How many different ways are there to pick up the triangle and place it back down inside its outline?

One can rotate the triangle over 120° around its symmetry axis (perpendicular to the plane and passing through its geometric center). Because the rotated triangle fits inside the outline, it looks identical in appearance to the initial triangle, making it impossible to tell whether an action has been carried out. The triangle is therefore said to be *symmetric*, or to possess a *geometric symmetry*. An alternative wording is to say the rotation over 120° is a *symmetry transformation* of the triangle and the triangle remains *invariant* under this symmetry operation.

There are other ways the triangle can be transformed without suffering any change in appearance. One could rotate the triangle over 240°, for instance, or reflect it

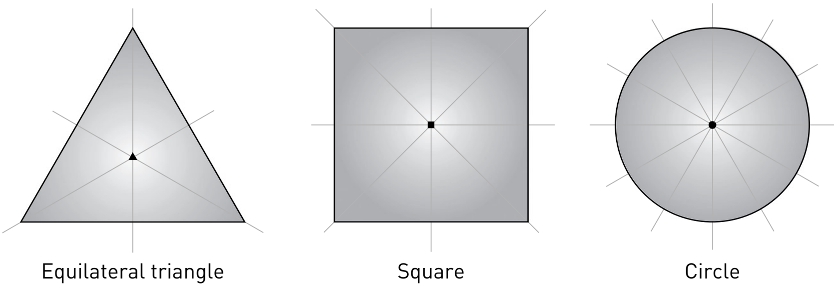


FIGURE 1.6 Three geometric figures with their reflection lines. Note that only six of the infinite number of reflection lines for the circle are actually shown.

---

[16] I. Stewart. *Why Beauty Is Truth: A History of Symmetry*. New York: Basic Books, 2007, pp. ix, 118.

[17] As Hermann Weyl once put it: "A thing is symmetrical if there is something you can do to it so that after you have finished doing it it looks the same as before." This definition was cited by physicist, Richard P. Feynman (1918–1988) in R. Feynman. *The Character of Physical Law*. Cambridge, MA: MIT Press, 1985, p. 84.

[18] *Equilateral* implies all sides are the same length.

within the plane through any of the reflection lines shown in Figure 1.6. You might be surprised to learn that doing nothing (i.e., rotating over 0° = 360°) is considered a valid symmetry transformation as well, termed the *identity transformation*. At first sight, it may seem a little pointless to include it in our list of symmetry operations, but as we will see later, we cannot do without it if we want to "speak group theory."

However, rotating the triangle over any other angle than the three listed so far results in a conformation that is *distinguishable* from the initial one. A rotation of 180°, for instance, would turn the triangle upside down, thus changing its overall appearance to an observer. Not *all* rotations are, therefore, *symmetry transformations* of the triangle, and in this case the triangle is said to be *asymmetric* with respect to these rotations.

In summary, the triangle possesses a total of six symmetries, six transformations that bring the triangle back into its original position: three rotations (over 0°, 120°, and 240°), and three reflections in the different mirror planes.

> As Bronowski said: At this point, the non-mathematician is entitled to ask: "So what? Is that what mathematics is about? Did Arab professors, do modern mathematicians, spend their time with that kind of elegant game?" To which the unexpected answer is—Well, it is not a game. It brings us face to face with something which is hard to remember, and that is that we live in a special kind of space—three-dimensional, flat—and the properties of that space are unbreakable. In asking what operations will turn a pattern into itself, we are discovering the invisible laws that govern our space. There are only certain kinds of symmetries which our space can support, not only in man-made patterns, but in the regularities which nature herself imposes on her fundamental, atomic structures.[19]

1.2.3 Quantifying symmetry

Consider the three geometric shapes in Figure 1.6. Which one, do you think, is the most symmetric? How can you tell?

With the modern definition of symmetry, we can easily calculate the *amount of symmetry* an object possesses; that is, we can *quantify* how symmetric an object is, and (when quantified) compare it with others on this basis. Basically, the more symmetries an object has (i.e., the more ways we can change the object without changing its overall appearance) the more symmetric we say it is.

Because doing nothing is a valid symmetry transformation, even the least symmetric of objects, such as a crumpled piece of paper or Einstein's haircut is considered symmetric. But clearly, an equilateral triangle is much more symmetric than Einstein's haircut; a square, in turn, is more symmetric than an equilateral triangle[20]; and a circle is much more symmetric than any of these objects.

Actually, a circle is one of the most symmetric objects around. Although a triangle and a square contain only a *finite* number of symmetry operations, a circle remains invariant under an *infinite* number of symmetry transformations. One can rotate a circle around its center through an infinite number of rotation angles, varying over

---

[19] J. Bronowski. *The Ascent of Man*. Boston: Little, Brown and Company, 1973, p. 174.

[20] A square has eight symmetries: four rotations over 0°, 90°, 180°, and 270°, and four reflections through the mirror lines shown in Figure 1.6.

the periodic interval $[0°, 360°]$, and reflect it in one of the infinite mirror axes passing through its geometric center (Figure 1.6).

1.2.4 Discrete and continuous symmetries

Figure 1.6 illustrates that there are two kinds of symmetries: either *continuous* or *discrete*.

Corollary 1.1 (Continuous and discrete symmetry): An object is said to have a *continuous* (*discrete*) symmetry when its symmetry transformations can be labeled by a set of *continuously* (*discretely*) varying parameters. ▪

The discreteness or continuity of symmetries is often dependent on the object concerned. A circle, for example, is said to have *continuous rotational symmetry* because its symmetry operations (i.e., rotations) can be labeled by a parameter (i.e., the rotation angle) that varies continuously over the interval $[0°, 360°]$. A triangle, on the other hand, has *discrete rotational symmetry* because the symmetry operations are labeled by a discretely varying parameter (i.e., by multiples of 120°). Some symmetries are always discrete (e.g., mirror reflections, permutations, and time reversals); others are always continuous (e.g., gauge symmetries in quantum field theory).

1.2.5 Multiplying symmetries

The study of *individual* symmetries can only lead us so far. It turns out to be much more interesting to look into the *complete set* of symmetries an object possesses. That is, to put the power of symmetry to its full use, we need to investigate how symmetries *interact* with one another, and determine which (mathematical) structure emerges from these interactions.

Consider, for the sake of illustration, the set of six symmetries of the equilateral triangle. One of the great insights of Galois involved *multiplying* symmetries—that is, performing one symmetry operation after another. Galois observed that by combining any two symmetry operations in pairs, you always get another symmetry from the set.

This should not be too surprising. After all, if R and S are two symmetry operations, both of which leave the triangle invariant, then clearly the successive application of these two symmetry operations leaves the triangle invariant as well. The two operations, R and S, carried out in succession must therefore be equivalent to another symmetry transformation, denoted SR.

For example, a rotation of the triangle through 120°, followed by another rotation through 240° (in counterclockwise fashion), is equivalent to a rotation through 360° (which is the identity transformation). Similarly, reflecting the triangle first about its vertical mirror axis and then rotating the triangle through 120°corresponds to a reflection of the triangle about its mirror axis through the lower right-hand vertex (Figure 1.7).

No matter how many symmetry operations you combine in this way, the end result will always be another symmetry transformation; if R and S are in the list, then so is SR. There is, in other words, no way of multiplying symmetries to create a *new* transformation that was not originally in the set of all symmetries; *you cannot jump out of the set.* This essential property of symmetry operations is called *closure.*

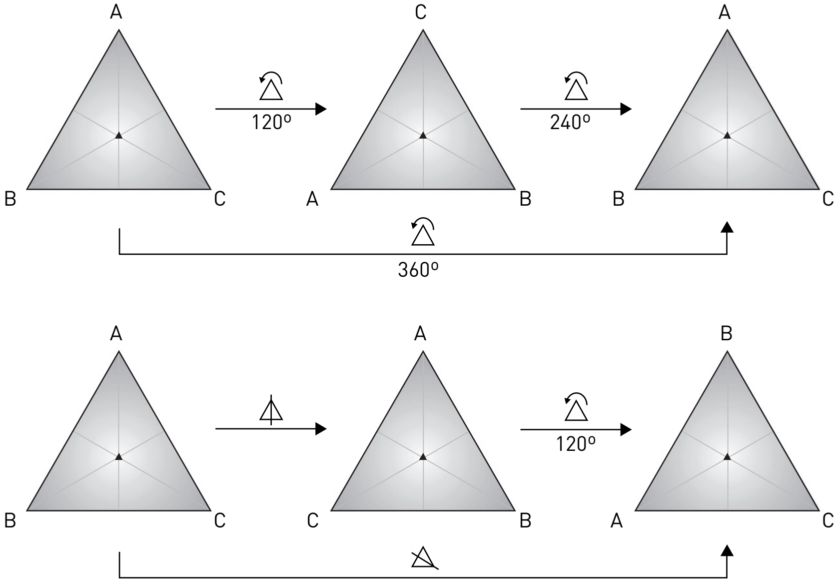


FIGURE 1.7 The vertices of the triangle are labeled *A*, *B*, and *C* to help you keep track of the movements. Ideally, however, these markings would not be present, making it utterly impossible to see whether a transformation has been carried out.

Closely related to this is the fact that every symmetry operation can be *undone* by performing another one. A counterclockwise rotation of 120°, for example, can be nullified by a clockwise rotation of 120°, which is equivalent to a counterclockwise rotation through 240°. In view of its inverting function, the rotation through 240° is therefore called the *inverse* of the 120° rotation. Analogously, because reflecting twice is the same as doing nothing, each reflection is considered to be its own inverse.

The notion of *inverse transformations*, along with the existence of an *identity transformation* and the fundamental property of *closure* leads us straight into the subject matter of group theory, the acquaintance of which we will make in the following chapter.

2 The elements of group theory

The Theory of Groups is a branch of mathematics in which one does something to something and then compares the result with the result obtained from doing the same thing to something else, or something else to the same thing.
–James R. Newman[1]

Galois was buried on Saturday, June 2, 1832. Thousands of friends, fellow Republicans, and students from the faculties of law and medicine attended the funeral procession. Galois had passed like a meteor,[2] barely twenty years old when his tragic life came to a sudden end. His complete mathematical oeuvre filled no more than one hundred pages. Yet, they contained the seeds of modern *group theory*, the subject matter of this chapter.

Chapter outline

The theory of groups is considered the language *par excellence* to study symmetry in science; it provides the mathematical formalism needed to tackle symmetry in a precise way. The aim of this chapter, therefore, is to lay the foundations of *abstract group theory*.

The central pillar on which these foundations rest, is the mathematical concept of a *group*, the algebraic structure of which will be defined formally in §2.1. In §2.2,

---

[1] J. R. Newman. “The Supreme Art of Abstraction: Group Theory”. In: *The World of Mathematics*, vol. 3. ed. J. R. Newman. New York: Simon and Schuster, 1956, p. 1534.

[2] G. Sarton. “Evariste Galois.” *Osiris* 3 (1937), p. 241–259.

an important distinction is made between *abstract groups* and *concrete groups*. This distinction serves as an explanation for the power and flexibility of group theory to handle highly diverse systems of sometimes completely different origins in much the same way.

This is followed by a discussion of *Abelian groups* in §2.3, which are exemplified in §2.4. In §2.5, the concept of *subgroups* is introduced, which naturally leads us to the idea of *symmetry breaking* in §2.6. The idea of broken symmetries will form a central theme throughout this book. Finally, §2.7 renders the ideas of §2.2 more explicit with a brief account of *isomorphisms* and *homomorphisms*.

What happened to the pile of papers Galois left on his desk, and how they paved the way for the development of the group concept during the nineteenth century, is the subject of §2.8, which closes this chapter. After these historical and philosophical diversions, and with the necessary definitions in place, we will have a powerful tool in hand to pursue our discussion of symmetry in Chapter 3, in which the rotational symmetry of the circle will be studied in greater detail.

2.1 MATHEMATICAL DEFINITION

The aim of this first section is to rephrase in more mathematical terms the rather dreamy definition of a group presented at the beginning of this chapter. Following James R. Newman (1907–1966), it will prove easiest to start from a *concrete group*, such as the symmetry group of the triangle (introduced at the end of Chapter 1), and proceed by stripping all the unessential details until one is left with the bare bones of the *abstract group* concept. Newman referred to this as the "*Carrollian method* of defining a grin as what remains after the Cheshire Cat, the vehicle of the grin, has vanished" (Figure 2.1).[3]

Let us thus replace the set of symmetry transformations of the triangle by a set of abstract elements, and the rule of performing one symmetry operation after another by an abstract combination rule. What is left after this process of abstracting is the notion of an *abstract group*—an algebraic structure consisting of a set of elements, together with an operation that combines any two of its elements to form a third element. To qualify as a group, the set of group elements and the operation must satisfy a few conditions that are called the *group axioms*.

Definition 2.1 (Group): Consider a nonempty *set* of distinct *elements*, $\mathcal{G} = \{e, a, b, c, d, \ldots\}$, endowed with a particular *law of composition* such that any two elements a and b of $\mathcal{G}$ can be combined to yield a third element, called the *product* of a and b and denoted $a \star b$ (or simply ab). The set $\mathcal{G}$ and the group multiplication operation $\star$ are said to form a *group* $(\mathcal{G}, \star)$ when they satisfy the following four requirements known as the *group axioms* or *group postulates*:

1. *Closure*. For all a, b in $\mathcal{G}$, the result of the operation $a \star b$ is also in $\mathcal{G}$. The set $\mathcal{G}$ is said to be *closed* (or *saturated*) under the law of composition.
2. *Associativity*. The law of composition is *associative*. That is, for all a, b, and c in $\mathcal{G}$, the *associative property* $(a \star b) \star c = a \star (b \star c)$ holds true. In other words, when you

---

[3] Newman, "The Supreme Art of Abstraction: Group Theory," pp. 1534–1535.

FIGURE 2.1 The *Cheshire Cat principle*. When group theory is applied to the natural world, the system under study is stripped from its physical content until an abstract entity remains, which can be handled by the machinery of abstract group theory. This is reminiscent of the Cheshire Cat, whose body vanishes slowly while Alice is looking at him. What remains is the Cat's weird grin—the group theoretical entity, which is treated as an abstraction of the physically real system. "Well! I've often seen a cat without a grin," said Alice; "but a grin without a cat! It's the most curious thing I ever saw in all my life!" Yet, the Cheshire Cat principle is what renders group theory, like so many other branches of mathematics, so powerful and diversely applicable in the real world.

evaluate a product of more than two elements, such as $a \star b \star c$, it doesn't matter how you group the elements, as long as you don't change their order. You can start by grouping a with b, making the product $a \star b = d$, and then form the product $d \star c$ corresponding to $(a \star b) \star c$. Alternatively, you could first combine b with c to obtain the product $b \star c = f$, and then make the composition $a \star f$, corresponding to $a \star (b \star c)$. Both ways of computing lead to the same result; hence, there is no ambiguity in writing $a \star b \star c$ and you are free to insert brackets wherever you want.

3. *Identity*. There exists a unique element e in $\mathcal{G}$ such that, for every element a in $\mathcal{G}$, the equation $e \star a = a \star e = a$ is obeyed. The element e is called the *unit element* or *identity element*.

4. *Invertibility*. For each a in $\mathcal{G}$, there exists an element a^{-1} in $\mathcal{G}$ such that $a \star a^{-1} = a^{-1} \star a = e$, where e represents the identity element and a^{-1} is called the *inverse element*. ▪

Corollary 2.1 (Finite and infinite groups): A group $(\mathcal{G}, \star)$ is said to be *finite* when it contains a finite number of group elements. This number is called the *order* of the group and is denoted by the symbols $o(\mathcal{G})$ or $|\mathcal{G}|$. *Infinite* groups contain an infinite number of elements and are said to be of infinite order. ▪

Finite groups are always discrete; infinite order groups, on the other hand, can be discrete, continuous, or a mixture of both.

Corollary 2.2 (Discrete and continuous groups): An infinite group $(\mathcal{G}, \star)$ is said to be *discrete* when its elements $e, a, b, c, \dots$ are *denumerably* infinite (that is, *countably* infinite), in the sense that they can be labeled by the natural numbers $1, 2, 3, \dots$. An infinite group $(\mathcal{G}, \star)$ is said to be *continuous* when its elements are *nondenumerably* infinite (that is, *noncountably* infinite); each element $e, a, b, c, \dots$ is then specified by a number n of continuously varying *parameters*, $r_1, r_2, \dots, r_n$. ▪

Upon one's first encounter with the definition of a group, one naturally feels somewhat confused, perhaps even a little sceptical. After all, it is far from clear *a priori* how much understanding, insight, and unification can follow from such simple axioms. One of the aims of this book is, therefore, to illustrate "what a wealth, what a grandeur of thought may spring from what slight beginnings."[4]

As the historian of mathematics Eric Temple Bell (1883–1960) famously wrote: "Whenever groups disclosed themselves [in the history of mathematical and scientific inquiry], simplicity crystallized out of comparative chaos."[5] The ubiquity of groups in areas as diverse as mathematics, physics, and chemistry has elevated the group concept to the central organizing principle of modern science, as we will happily show in the chapters to follow.

2.2 THE ABSTRACT AND THE CONCRETE

Much of the power of group theory lies in its complete detachment from anything concrete. The theory of groups has, unsurprisingly, been called "the supreme art of mathematical abstraction."[6] Indeed, in the formal definition of a group, we consciously refrained from specifying the exact nature of the group elements and their product rule. The elements $e, a, b, c, \dots$ were introduced as *abstract* entities, forming *abstract* groups under *abstract* combination laws. By focusing in this way on the grin, rather than on the entire Cheshire Cat (Figure 2.1), group theory concerns itself only with "the fine filigree of underlying relationships";[7] it studies the relations regardless

---

[4] British geometer Henry Frederick Baker (1866–1956), as quoted by American historian of mathematics Florian Cajori (1859–1930) in F. Cajori. *A History of Mathematics*. 2nd ed. New York: MacMillan, 1919, p. 283.

[5] E. T. Bell. *Mathematics, Queen and Servant of Science*. New York: McGraw-Hill, 1951, p. 164.

[6] Newman, "The Supreme Art of Abstraction: Group Theory," p. 1534.

[7] Ibid.

FIGURE 2.2 Having studied the *group theoretical* grin of the system under study, one can make its *physical* body reappear.

of the nature of the relata. It is this ignoring of all the unessential details that has turned group theory into "the most powerful instrument yet invented for illuminating structure."[8]

Of course, at some point, the cat's body has to reappear (Figure 2.2). That is, for group theory to be useful in science, the *abstract* elements and product rules have to be made *concrete*; they have to acquire specific meanings by going into correspondence with their concrete counterparts. "Otherwise, manipulating the group amounts to nothing more than a game, and a pretty vague and arid game at that, suitable only for the most withdrawn lunatics."[9] As Cassius J. Keyser (1862–1947) observed:

> A great idea is always generic and abstract but it has its living significance in the particular and concrete—in a countless multitude of differing instances or examples of it.[10]

Abstract groups can be *represented* (or *realized*) by an infinite number of concrete groups. The exact relationship between an abstract and a concrete group (or between two or more concrete groups) will be made more explicit in §2.7. For the moment, let us note only that these concrete elements are often *arithmetic objects*, such as numbers, vectors, or matrices; they can also be *mathematical operations* or *physical operations*, such as rotations and other symmetry transformations. Most of our attention in this book will be focused on groups with elements that are symmetry transformations.

In much the same way, the abstract product symbol $\star$ can be realized concretely to denote just about *any* kind of combination operation. It doesn't merely symbolize the arithmetic operation of *multiplying* two group elements ($a \times b$) as its (misleading) name may seem to imply at first sight; it could just as well represent a simple *addition* of the group elements ($a + b$), or it might be used to denote the *consecutive application* of two symmetry transformations.

---

[8] Newman, "The Supreme Art of Abstraction: Group Theory," p. 1534.

[9] Ibid., p. 1535.

[10] C. J. Keyser. *Mathematical Philosophy: A Study of Fate and Freedom*. New York: Dutton, 1922, p. 205.

In the case of rotation operators, for instance, the product of two rotations over the respective angles ω_2 and ω_1, written as $R(\omega_2) \star R(\omega_1)$, simply means to carry out both rotations consecutively, first over ω_1 and subsequently over ω_2. Notice that the operations are *right justified*—meaning that the sequence of operations should be applied from right to left (and *not* from left to right, although this is the traditional direction for reading and writing in English). Think of an object or "target" on which the operators are acting that is placed behind the operators. Then, the first operator that "hits" the target is, indeed, $R(\omega_1)$ and the second is $R(\omega_2)$. The same rule applies to differential operators; for example,

$$\frac{d}{dx} x f(x) \tag{2.1}$$

means *first* to multiply the function $f(x)$ by x, and, *secondly*, to differentiate the result. This is, of course, completely different from

$$x \frac{d}{dx} f(x), \tag{2.2}$$

where the function $f(x)$ is first differentiated and then multiplied by x.

2.3 ABELIAN GROUPS

The order in which the group operations are carried out is therefore significant; the combination of element a with element b is not necessarily the same as the combination of element b with element a; the equation $a \star b = b \star a$ does not always hold (Figure 2.3). Whenever the equation *does* hold true, the elements a and b are said to *commute*. It follows from Definition 2.1 that the unit element e commutes with all group elements. Similarly, every element a commutes with its

FIGURE 2.3 A mad tea party. " 'You should say what you mean,' the March Hare went on. 'I do,' Alice hastily replied; 'at least—at least I mean what I say—that's the same thing, you know.' 'Not the same thing a bit!' said the Hatter. 'You might just as well say that "I see what I eat" is the same thing as "I eat what I see!" ' 'You might just as well say,' added the March Hare, 'that "I like what I get" is the same thing as "I get what I like!" ' 'You might just as well say,' added the Dormouse, who seemed to be talking in his sleep, 'that "I breathe when I sleep" is the same thing as "I sleep when I breathe!" ' 'It *is* the same thing with you,' said the Hatter, and here the conversation dropped."

inverse element a^{-1}. And of course, every element also commutes with itself. Groups for which *all* the elements commute with one another are called *Abelian* in honour of Norwegian mathematician Niels Hendrik Abel (1802–1829), and the product rule is said to be *commutative*.

Definition 2.2 (Abelian groups): A group $(\mathcal{G},\star)$ is said to be *Abelian* if all the group elements *commute* (i.e., if $a \star b = b \star a$ for every pair of elements $a,\ b \in \mathcal{G}$). The multiplication rule $(\star)$ is then said to be *commutative*. Conversely, all groups with a *noncommutative* product rule are called *non-Abelian*. ▪

As an example of a concrete Abelian group, consider the set of all positive and negative integers, including 0 and denoted $\mathbb{Z} = \{0,\ \pm 1,\ \pm 2,\ \pm 3,\ \ldots\}$. This set forms an *integer addition group* $(\mathbb{Z},+)$ under the mathematical operation of addition. After all, the set is clearly closed under addition and the law of associativity always holds:

$$(a+b)+c = a+(b+c). \tag{2.3}$$

The number 0 represents the identity element since

$$0+a = a+0 = a, \tag{2.4}$$

and $-a$ is the inverse element of a because

$$a+(-a) = (-a)+a = 0. \tag{2.5}$$

This fulfils the four group axioms. Now, because of the *commutativity of addition*, the equation $a+b = b+a$ holds true for all integers $a, b \in \mathbb{Z}$; this proves the Abelian character of the group $(\mathbb{Z},+)$. Also, $\mathbb{Z}$ contains an infinite number of elements, and because one cannot proceed *continuously* from one element to the other, $(\mathbb{Z},+)$ is a discrete Abelian group of infinite order.

Not all groups are Abelian, however; most symmetry groups, for instance, are non-Abelian, as we will see in the chapters to follow.

2.4 EXAMPLES OF GROUPS

We now consider some simple examples of groups and represent them in tabular form.

1. The smallest finite group $\{e\}$ consists of the identity element e alone and is of order 1, with $e \star e = e$. This is also called the *trivial group*.
2. The set of real numbers $\mathbb{R}$ forms a group $(\mathbb{R},+)$ under the operation of addition. This group is, in many aspects, comparable with the integer addition group $(\mathbb{Z},+)$; the identity element is 0 and the inverse of a is $-a$. Also, because addition in $\mathbb{R}$ is commutative, the group $(\mathbb{R},+)$ is Abelian. However, the set of real numbers under addition is a typical example of a *continuous* one-parameter group, unlike the *discrete* addition group $(\mathbb{Z},+)$. After all, the elements of $\mathbb{R}$ are *nondenumerable* and the parameter of the group is just the real number itself, which varies continuously from $-\infty$ to $+\infty$.
3. The set of complex numbers $\mathcal{G} = \{1,\ i,\ -1,\ -i\}$ forms a finite group $(\mathcal{G},\times)$ of order 4 with respect to the multiplication operation. This can be seen clearly by constructing a so-called *multiplication table* in Table 2.1:

Table 2.1 Multiplication table for the group $\{1, i, -1, -i\}$

	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

Table 2.2 Example of a multiplication table. The product $a \star b$ is found at the intersection of the row starting with a and the column headed by b

	e	a	$\cdots$	b	$\cdots$
e	$e \star e$	$e \star a$	$\cdots$	$e \star b$	$\cdots$
a	$a \star e$	$a \star a$	$\cdots$	$a \star b$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	
b	$b \star e$	$b \star a$	$\cdots$	$b \star b$	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$

Known variously as a *composition table* or a *Cayley table* (after British mathematician Arthur Cayley [1821–1895]), a multiplication table portrays the structure of finite groups. To construct the table, start by listing all four group elements in the top row (in no particular order, but without repeating elements), then do the same in the leftmost column (again, in no particular order). This traces out the boundaries of a 4 × 4 square table. Each entry in the grid is then obtained by making the cross product between an element from the leftmost column and one from the top row (Table 2.2). Because many groups are non-Abelian, the convention of first writing the element from the leftmost column and then writing the element from the top row is of the utmost importance. A closer examination of the multiplication table verifies the four group axioms:

I. Since all entries in the table are elements of $\mathcal{G}$, the set $\mathcal{G}$ is closed under the operation of multiplication.

II. All elements of $\mathcal{G}$ are complex numbers: $\mathcal{G} = \{1,\ i,\ -1,\ -i\} \subset \mathbb{C}$. Since multiplication in $\mathbb{C}$ is associative, the associative law is obeyed in $(\mathcal{G}, \times)$.

III. The column headed by the element 1 of the top row matches the leftmost column—meaning, that: $1 \times 1 = 1$, $i \times 1 = i$, $-1 \times 1 = -1$, and $-i \times 1 = -i$. In short, $a \times 1 = a \; \forall a \in \mathcal{G}$. Analogously, the row headed by the element 1 is identical to the top row: $1 \times (1) = 1$, $1 \times (i) = i$, $1 \times (-1) = -1$, and $1 \times (-i) = -i$, which can be summarized as $1 \times a = a \; \forall a \in \mathcal{G}$. Bringing together these two observations, we have just verified that $1 \times a = a \times 1 = a \; \forall a \in \mathcal{G}$, thus identifying 1 as the unit element of $(\mathcal{G}, \times)$.

IV. For every element in the top row there is a corresponding element in the leftmost column (and vice versa) whose product equals the identity element 1. In other words, each element of $\mathcal{G}$ possesses a unique inverse element. The identity element, by definition, is its own inverse since $1 \times 1 = 1$. The same holds true for the element -1 because $(-1) \times (-1) = 1$. Finally, i and $-i$ are each other's inverse: $i \times (-i) = (-i) \times i = 1$.

Since multiplication in $\mathbb{C}$ is commutative, the group $(\mathcal{G}, \times)$ is Abelian. This fact can also be deduced from the multiplication table. Because the ordering of the group elements in the top row (left to right) matches the ordering in the leftmost column (up to down), the multiplication table is *symmetric* along its *diagonal axis* (i.e., $a \times b = b \times a \, \forall a, b \in \mathcal{G}$).

4. As a final example, consider the symmetry group of the equilateral triangle. Let e denote the null operation; p and q signify the rotations through 120° and 240°, respectively; and r, s, and t represent the three reflections through the lines r, s, and t, as shown in Figure 2.4. Now, try constructing the multiplication table yourself. It may be helpful to cut an equilateral triangle out of paper or cardboard and to label its vertices A, B, and C. You can then play with the triangle by performing one operation after the other and writing down your results, which should look like Table 2.3[11]: This group is clearly closed and associative. The identity element is e, and each element has an inverse; e, r, s, and t are their own inverse whereas p and q are each other's inverse.

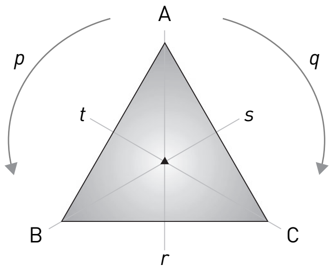


FIGURE 2.4 The six symmetry transformations of an equilateral triangle. The identity transformation e is not indicated.

Table 2.3 Multiplication table for the symmetry group of the equilateral triangle

	e	p	q	r	s	t
e	e	p	q	r	s	t
p	p	q	e	t	r	s
q	q	e	p	s	t	r
r	r	s	t	e	p	q
s	s	t	r	q	e	p
t	t	r	s	p	q	e

---

[11] Reflections may be executed effectively by pivoting the triangles around the dashed lines in Figure 2.4. Note that the symmetry operation changes the position of the vertices. However, the orientation of the symmetry planes are left immobile in space. See A. Ceulemans. *Group Theory Applied to Chemistry*. Dordrecht: Springer, 2013, p. 23.

Notice also that this group is non-Abelian; for instance, $t \star r = p$, but $r \star t = q$. As a consequence, the Cayley table is not symmetric along its main diagonal.

2.5 SUBGROUPS

To introduce the concept of a *subgroup*, imagine restricting the set of symmetry transformations $\mathcal{G} = \{e, p, q, r, s, t\}$ of the equilateral triangle to rotations only. Interestingly, this new set $\mathcal{H} = \{e, p, q\}$ forms a group as well, as can be seen from the multiplication table. Since $\mathcal{H}$ is a *subset* of $\mathcal{G}$, forming a group $(\mathcal{H},\star)$ under the same composition law as $(\mathcal{G},\star)$, $(\mathcal{H},\star)$ is called a *subgroup* of $(\mathcal{G},\star)$.

Not every subset of $\mathcal{G}$ forms a subgroup, however. The set of reflections $\mathcal{I} = \{r, s, t\}$, for instance, does *not* yield a subgroup of $(\mathcal{G},\star)$, since $(\mathcal{I},\star)$ is not closed. For example, $s \star t = p$, but p is not an element of set $\mathcal{I}$. Set $\mathcal{I}$ also lacks an identity element e. All of this is easily deduced from the previous multiplication table. Generalizing:

Definition 2.3 (Subgroup): If, from group $(\mathcal{G},\star)$ a new group $(\mathcal{H},\star)$ can be formed under the same composition law $\star$ by taking a subset $\mathcal{H}$ of $\mathcal{G}$, denoted $\mathcal{H} \subset \mathcal{G}$, then $(\mathcal{H},\star)$ is said to form a *subgroup* of $(\mathcal{G},\star)$. ■

Every group $(\mathcal{G},\star)$ is a subgroup of itself. Also, the identity element e always forms a *trivial subgroup* $\{e\}$ of $(\mathcal{G},\star)$ of order 1. These are called *improper* subgroups. Subset $\mathcal{H} = \{e, p, q\}$ of $\mathcal{G} = \{e, p, q, r, s, t\}$, on the other hand, is said to form a *proper* subgroup $(\mathcal{H},\star)$ of group $(\mathcal{G},\star)$.

2.6 SYMMETRY BREAKING

The subgroups of a group are usually denoted as a *chain of subgroups*; the parent group is listed first, followed by the different subgroups in decreasing order:

$$(\mathcal{G},\star) \supset (\mathcal{G}_1,\star) \supset (\mathcal{G}_2,\star) \supset (\mathcal{G}_3,\star) \supset \ldots, \tag{2.6}$$

with $\mathcal{G}_1$ a subset of $\mathcal{G}$, $\mathcal{G}_2$ a subset of $\mathcal{G}_1$, and $\mathcal{G}_3$ a subset of $\mathcal{G}_2$. In the case of the symmetry group of the equilateral triangle, for example,

$$(\mathcal{G},\star) \supset (\mathcal{H},\star) \supset \{e\}. \tag{2.7}$$

This chain actually represents a series of *symmetry breakings*. In the first step, the full symmetry of the triangle is broken because reflections are no longer included. In the second step, the remaining rotational symmetry is broken even further to yield the trivial group, where the only symmetry left is the null operation.

2.7 ISOMORPHISMS AND HOMOMORPHISMS

Let us return, for a moment, to our discussion of abstract and concrete groups. Recall that when the notion of an abstract group was outlined in Definition 2.1, no direct allusion was made to the *nature* of its elements. We thus emphasized in §2.2 that, within *abstract group theory*, the elements of a group are just meaningless *symbols*, devoid of any physical reality, that combine according to a given product rule, the structural properties of which are summarized in a multiplication table. By way of illustration, consider the following set of abstract elements $\mathcal{G} = \{e, a, b, c\}$, forming

Table 2.4 Multiplication table of an abstract group of 4 elements

	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

an abstract group $(\mathcal{G},\star)$ of finite order 4 with multiplication table as given in Table 2.4. By inspection of the multiplication table, we can easily verify that all four group axioms are satisfied, which proves that $(\mathcal{G},\star)$ is, indeed, a group.

We then were able to see that whenever the nature of the elements and the product rule are specified, a *concrete group* is obtained, which is just one out of an infinite number of possible *realizations* of the *abstract group*; one says the abstract group has been *represented* by the concrete group. The aim of this section is to outline how such representations can be realized.

Basically, every abstract element has to be brought into correspondence with a concrete element. This is done via a *mapping* ϕ from the abstract set $\mathcal{G}$ to the concrete one $\mathcal{G}'$, which we denote as

$$\mathcal{G} \xrightarrow{\phi} \mathcal{G}' \quad \text{or} \quad \phi : \mathcal{G} \to \mathcal{G}'. \tag{2.8}$$

Consider, for example, the following mapping from the elements of $\mathcal{G} = \{e,\ a,\ b,\ c\}$ to the elements of $\mathcal{G}' = \{1,\ i,\ -1,\ -i\}$:

$$\begin{array}{ccr} \mathcal{G} & \to & \mathcal{G}' \\ \hline e & \mapsto & 1 \\ a & \mapsto & i \\ b & \mapsto & -1 \\ c & \mapsto & -i \end{array} \tag{2.9}$$

Each element g of $\mathcal{G}$ corresponds to an element g' of $\mathcal{G}'$, which is called the *image* of g. This mapping is represented as $g \mapsto g'$. Thus, 1 is the image of e, i is the image of a, -1 is the image of b, and $-i$ is the image of c. Let us assume the abstract product $\star$ is realized by the multiplication operation $\times$. Notice, then, that this mapping preserves the composition rule; for instance,

$$b \star c = a \quad \overset{\text{is mapped to}}{\longleftrightarrow} \quad (-1) \times (-i) = i. \tag{2.10}$$

The abstract group $(\mathcal{G},\star)$ is therefore said to be represented by the concrete group $\left(\mathcal{G}',\times\right)$, the multiplication table of which was already given on page 22.

Because there is a *one-to-one correspondence* between the elements of $\mathcal{G}$ and $\mathcal{G}'$ (i.e., each element of $\mathcal{G}'$ is the image of one and only one element of $\mathcal{G}$) both groups are of the same order and their multiplication tables are equivalent (differing only in the designation of the elements). Groups that share the same algebraic structure are said to be *isomorphic* to one another, and the mapping is called an *isomorphism.*

Another isomorphic realization of the abstract group $(\mathcal{G},\star)$ can be obtained by considering the rotational symmetry of a square. Since a square is invariant under rotations through multiples of 90°, we obtain the following set of symmetry

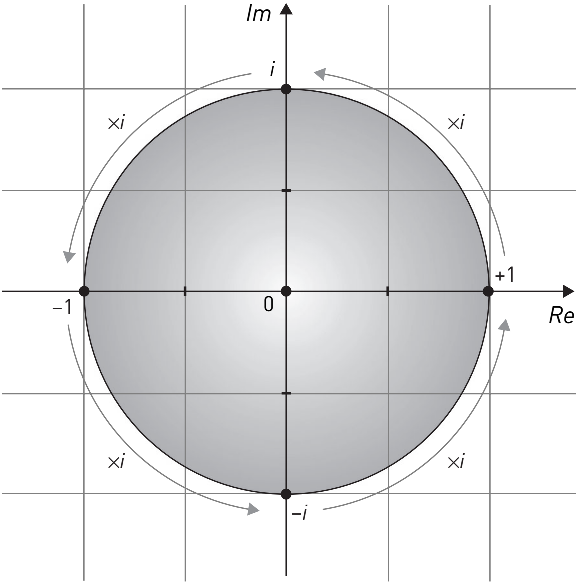


FIGURE 2.5 The isomorphism between the two concrete groups $(\mathcal{G}', \times)$ and $(\mathcal{G}'', \star)$ is particularly clear in the *Argand plane*,[12] where multiplication by i is seen to amount to a counterclockwise turn of 90°.

transformations: $\mathcal{G}'' = \{0^\circ, 90^\circ, 180^\circ, 270^\circ\}$. We can then make the following mapping from $\mathcal{G}$ to $\mathcal{G}''$:

$$\begin{array}{ccc} \mathcal{G} & \to & \mathcal{G}'' \\ \hline e & \mapsto & 0^\circ \\ a & \mapsto & 90^\circ \\ b & \mapsto & 180^\circ \\ c & \mapsto & 270^\circ \end{array} \tag{2.11}$$

This mapping preserves the composition rule, so $(\mathcal{G}'', \star)$ is another concrete realization of $(\mathcal{G}, \star)$. And because the mapping is *one to one*, the symmetry group $(\mathcal{G}'', \star)$ is isomorphic to $(\mathcal{G}, \star)$. Moreover, this implies that the two concrete groups $(\mathcal{G}', \times)$ and $(\mathcal{G}'', \star)$ are isomorphic to one another under the mapping: $1 \mapsto 0^\circ$, $i \mapsto 90^\circ$, $-1 \mapsto 180^\circ$, and $-i \mapsto 270^\circ$ (Figure 2.5).

What this actually means is that we can limit ourselves to a study of the abstract group $(\mathcal{G}, \star)$ and then translate our findings according to the mappings in Eqs. (2.9) and (2.11) to yield valid information about both concrete groups $(\mathcal{G}', \times)$ and $(\mathcal{G}'', \star)$. For example, the finding that "a combines with b to yield c" translates to "the complex number i multiplied by -1 yields the complex number $-i$," or "a rotation of the square by 180°, followed by another one through 90°, is the same as one rotation through 270°." Here is a first example of the power of group theory at work.

---

[12] Complex numbers are presented in a plane formed by the real axis and the orthogonal imaginary axis. It is named after the French mathematician Jean-Robert Argand (1768–1822).

Table 2.5 Multiplication tables corresponding to a many-to-one mapping

	e	b	a	c
e	e	b	a	c
b	b	e	c	a
a	a	c	b	e
c	c	a	e	b

$\leftrightarrow$

	1	-1
1	1	-1
-1	-1	1

We will encounter many isomorphisms in this book. The symmetries of elementary particle physics, for instance, will reveal themselves to be isomorphic to the symmetries of the harmonic oscillator (Chapter 7). Similarly, the symmetries of the hydrogen atom will be shown to be isomorphic to the symmetries of the solar system (Chapter 9).

Now, before closing this section, consider yet another mapping between the elements of $\mathcal{G} = \{e,\ a,\ b,\ c\}$ and the set $\mathcal{G}''' = \{1,\ -1\}$:

$$\begin{array}{ccr} \mathcal{G} & \rightarrow & \mathcal{G}''' \\ \hline e & \mapsto & 1 \\ a & \mapsto & -1 \\ b & \mapsto & 1 \\ c & \mapsto & -1 \end{array} \tag{2.12}$$

Group multiplication is still preserved; $b \star c = a$, for instance, gets mapped onto $1 \times (-1) = -1$. So, $\left(\mathcal{G}''', \times\right)$ is another representation of the abstract group $(\mathcal{G}, \star)$. However, because there is a *many-to-one correspondence* between both sets of elements (i.e., different elements from $\mathcal{G}$ are mapped onto one and the same element from $\mathcal{G}'''$), the orders of the groups differ: $|\mathcal{G}| \neq \left|\mathcal{G}'''\right|$. The algebraic structure of both groups is therefore different, as can be seen by inspection of both multiplication tables (Table 2.5). Groups $(\mathcal{G}, \star)$ and $\left(\mathcal{G}''', \times\right)$ are said to be *homomorphic* to one another, and the mapping is called a *homomorphism*. Generalizing the aforementioned:

Definition 2.4 (Isomorphisms and homomorphisms): A *mapping* ϕ of $\mathcal{G}$ onto $\mathcal{G}'$ allows each element a of $\mathcal{G}$ to be assigned to an element $a' = \phi(a)$ of $\mathcal{G}'$. If ϕ is a *one-to-one* mapping between the elements of $\mathcal{G}$ and $\mathcal{G}'$ for which the composition rule is preserved:

$$a \star b = ab \quad \longleftrightarrow \quad \phi(a) \star' \phi(b) = \phi(ab), \tag{2.13}$$

then ϕ is called an *isomorphism* and $\mathcal{G}$ and $\mathcal{G}'$ are said to be *isomorphic* to one another: $\mathcal{G} \approx \mathcal{G}'$, with $|\mathcal{G}| = \left|\mathcal{G}'\right|$. When ϕ is a *many-to-one* mapping that preserves the product rule, then ϕ is called a *homomorphism* and $\mathcal{G}$ and $\mathcal{G}'$ are said to be *homomorphic* to one another, with $|\mathcal{G}| \neq \left|\mathcal{G}'\right|$. ■

2.8 HISTORICAL INTERLUDE

In this final section, the development of group theory throughout the nineteenth century is briefly described. Who were the first explorers of the group theoretical concept? Which mathematicians first set foot on the group theoretical shores of the

Platonic world? And who was the first mathematician to write down explicitly the four group axioms?

Just as with the concept of symmetry, the development of the abstract group concept was a remarkably slow process. What is more, group theory did not emerge from a study of *geometric symmetries*, but from the study of *algebraic symmetries* that occur in equation theory.[13]

2.8.1 Évariste Galois

According to recorded history, Galois was probably the first to explicitly define the notion of a group in 1832. Of course, mathematicians had been using groups well before the first formal definitions appeared in print. As we saw in §2.4, most number systems (e.g., the natural, rational, real, and complex numbers) have a group structure. The same can be said about rotations, translations, or even modular arithmetic. So, whoever was studying these matters was (at least unconsciously) dealing with groups.

But let us return to Galois. When Galois first composed his *Premier Mémoire* back in 1829, he failed to provide a definition of the term *group*, and readers had to guess its technical meaning from Galois' repeated use of it in the document. Therefore, it is not really surprising that Poisson deemed Galois' paper to be "incomprehensible" (see Chapter 1). It was only on the eve of the duel, when revising his manuscript for the final time, that Galois added some last-minute marginal additions in which he wrote down the first technical definition of a group. He thus wrote: "if in such a group one has the substitutions S and T, one is sure to have the substitution ST."[14]

Galois' brother Alfred and his good friend Auguste Chevalier took upon themselves the arduous task of collecting and cataloguing Galois' mathematical manuscripts. His *Lettre Testamentaire* appeared in the September issue of the *Revue Encyclopédique*, along with a short obituary from the hand of Chevalier.[15] And during the summer of 1843, Chevalier donated all the manuscripts to mathematician Joseph Liouville (1809–1882). Profoundly impressed by Galois' revolutionary findings, Liouville promised the Paris Academy to publish Galois' complete *Œuvres Mathématiques*, which he finally did in his journal in 1846—fourteen years after Galois' untimely death.[16]

---

[13] For this reason, the term *group* has been long synonymous with *permutation group*.

[14] Translated from French by Peter Neumann in P. M. Neumann. *The Mathematical Writings of Évariste Galois*. Heritage of European Mathematics. Zürich: European Mathematical Society Publishing House, 2011, p. 115. The original text reads: "si dans un pareil groupe on a les substitutions S et T, on est sûr d'avoir la substitution ST."

[15] E. Galois. "Travaux Mathématiques d'Évariste Galois: Lettre de Galois." *Revue Encyclopédique* 55 (1832), pp. 568–576; A. Chevalier. "Nécrologie Evariste Galois." *Revue Encyclopédique* 55 (1832), pp. 744–754.

[16] E. Galois. "Oeuvres Mathématiques d'Évariste Galois." *Journal de Mathématiques Pures et Appliquées* XI (1846), pp. 381–444. At that time, Liouville's journal was also known as the *Journal de Liouville*. Galois' *Œuvres Mathématiques* was later reprinted in book form in E. Picard, ed. *Œuvres Mathématiques d'Évariste Galois*. Paris: Gauthier-Villars, 1897. For the English translation of Galois' œuvre, see Neumann, *op. cit.*

2.8.2 The French school

Interestingly, Augustin-Louis Cauchy came up with a very similar definition in 1845, the year before Galois' publications came into print in Liouville's journal. Cauchy did not adopt the term *group*, however, but preferred the phrase *conjugate system of substitutions*.[17] One might wonder to what extent Cauchy had been influenced by the reading of Galois' *Premier Mémoire*, which he submitted to the Academy in 1829.

In any case, both Galois and Cauchy defined a group in terms of the closure property. There was no explicit statement of the associative law or the existence of an identity element and inverse elements. But, judging from their work, these were clearly assumed implicitly.

Meanwhile, Liouville had started lecturing on Galois' work. Among those who attended the course were Joseph Serret (1819–1885), Joseph Bertrand (1822–1900), and Charles Hermite (1822–1901). All of them would later contribute to the subject, but it was Camille Jordan (1838–1922), a student of Serret, who properly started the systematic study of finite groups. Having busied himself from 1860 till 1870 with the theory of groups, Jordan's research culminated in the publication of his *Traité des Substitutions et des Équations Algébriques* in 1870, the first major textbook on group theory, which helped to bring Galois' theory into the spotlight.

2.8.3 Sir Arthur Cayley

The first attempt at an *abstract* definition of a group arose in the work of the British mathematician Sir Arthur Cayley (1821–1895), who wrote two papers in 1854.[18] Cayley wrote:

> A set of symbols, $1, \alpha, \beta, \ldots$ all of them different, and such that the product of any two of them (no matter in what order), or the product of any one of them into itself, belongs to the set, is said to be a *group*. ... These symbols are not in general convertible (commutative) but are associative. ... It follows that if the entire group is multiplied by any one of the symbols, either as further or nearer factor [i.e., left or right multiplication], the effect is simply to reproduce the group.[19]

Notice that Cayley refers to a group as a set of elements (symbols) with a binary operation (product). The closure property and associative law are mentioned explicitly, and the existence of an identity and inverse elements follows implicitly from the properties of left and right cancelation. In the same article, Cayley also

---

[17] Both names were used interchangeably until 1880. After the appearance of Camille Jordan's commentary on Galois' work, the word *group* came to be adopted as the standard term. This was reinforced only with Jordan's publication of his *Traité des Substitutions et des Équations Algebraique* in 1870.

[18] A. Cayley. "On the Theory of Groups, As Depending on the Symbolic Equation $\theta^n = 1$." *Philosophical Magazine* VII (1854), pp. 40–47, 408–409.

[19] Cayley, "On the Theory of Groups, As Depending on the Symbolic Equation $n = 1$," p. 41. In a footnote to this definition, Cayley acknowledged that "the idea of a group as applied to permutations or substitutions is due to Galois, and the introduction of it may be considered as marking an epoch in the progress of the theory of algebraical equations."

introduced the so-called *Cayley table* of a group (see §2.4), and realized that matrices and quaternions could form groups as well.

Cayley's papers were well ahead of their time, and they failed to make a lasting impression on his fellow mathematicians. However, when Cayley returned to the subject in 1878,[20] the time was finally ripe for group theory to become a subject worthy of deeper consideration.

---

[20] A. Cayley. "A Theorem on Groups," *Mathematische Annalen* XIII (1878), pp. 561–565; A. Cayley. "On the Theory of Groups," *Proceedings of the London Mathematical Society* IX (1878), pp. 126–133; A. Cayley. "The Theory of Groups," *American Journal of Mathematics* I (1878), pp. 50–52; A. Cayley. "The Theory of Groups: Graphical Representation." *American Journal of Mathematics* I (1878), pp. 174–176.

3 The axial rotation group

Μή μου τοὺς κύκλους τάραττε. *[Do not disturb my circles!]*
–Last words of Archimedes of Syracuse (278–212 BC)[1]

If the first chapter of this book revolved around the *concept of symmetry* and the second chapter around the *concept of a group*, then this chapter begins the *grand synthesis* of both. That's right, the relevance of the group concept for physics and chemistry has its origin in the fundamental connection between (1) the algebraic structure of a group and (2) the symmetry transformations of the particular system under study. That is to say, the complete *set* of symmetry transformations of an object along with the *operation* of combining two such transformations by performing one after the other forms a *symmetry group*, as we shall see in this chapter.

Although most discussions of group theory start with a consideration of finite, *discrete symmetries*, such as the symmetries of the triangle (§1.2.2), this monograph plunges immediately into the fascinating but often perplexing world of *continuous symmetries* that lie at the foundations of both chemistry and physics. Let's embark on a spectacular journey toward the sublime symmetries of both atomic and molecular chemistry, nuclear and elementary particle physics. Such a journey requires

---

[1] According to legend, ancient Greek mathematician and astronomer Archimedes of Syracuse was drawing circles in the sand when a Roman soldier disturbed him. With his sword drawn over Archimedes' head, the soldier asked him to reveal his identity, but Archimedes' eyes and mind were fixed on the mathematical problem in front of him, and he only uttered: "Do not disturb my circles!" Confused by this answer and without respect for the art of mathematics, the soldier killed Archimedes.

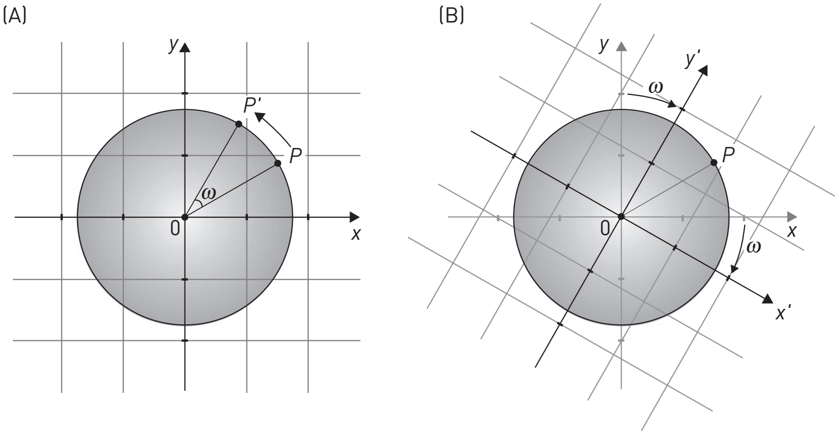
careful preparation, though, and we first need to assemble the necessary mathematical tools before we can tackle the group theoretical foundations of these sciences in Part II of this book.

Chapter outline

In this chapter, we therefore consider the symmetry of the circle—icon of perfection—a bit more carefully than before. After making a distinction between the *active* and *passive* ways of defining symmetry in §3.1, we will examine the different symmetry transformations of the circle by *rotation operators* in §3.2. We then take a look at how these rotation operators form a symmetry group—the *axial rotation group*—in §3.3. In the next two sections, §§3.4 and 3.5, we investigate how *coordinates* and *coordinate functions* transform under these rotation operators by expressing the two-dimensional rotations of the circle algebraically by means of *rotation matrices*. This will lead us to consider the concept of *matrix representations* in §3.6. Finally, in §3.7, we will prove that the axial rotation group is *isomorphic* to the so-called *special orthogonal group* in two dimensions, denoted SO(2) for short (with “S” for special, “O” for orthogonal, and “2” for two dimensions).

3.1 ACTIVE VERSUS PASSIVE VIEW OF SYMMETRY

Let’s take off by drawing a circle in the two-dimensional Euclidean plane $\mathbb{R}^2$, centred at the origin O of a Cartesian coordinate frame $\Sigma(x, y)$ (Figure 3.1). The origin O functions as the fixed point (or pivot point) of rotation. There are two ways to express the rotational symmetry of the circle:



Active transformation

Passive transformation

FIGURE 3.1 (A) In the *active transformation*, point P is rotated actively in a counterclockwise fashion about the pivot point O over an angle ω to the new point P'. (B) In the *passive transformation*, point P does not move, but the coordinate system $\Sigma(x, y)$ is rotated in clockwise fashion about the origin over an angle ω to the new coordinate system $\Sigma'(x', y')$. Now, P' is to $\Sigma(x, y)$ what P is to $\Sigma'(x', y')$.

1. We can rotate the circle around its symmetry axis through a particular angle ω, and notice that it remains invariant under this rotation.
2. Alternatively, we can also keep the circle fixed while changing our perspective of it by tilting our head a little to the left or to the right. If you imagine the Cartesian frame to be attached to your face, this actually corresponds to a rotation of the coordinate system from which the circle is observed.[2] Once again, the appearance of the circle remains unchanged during this change of perspective.

Given that, in the first way, every point of the circle changes position during the transformation, it is referred to as an *active transformation*. In the second way, the circle is kept in a fixed position. So, from the circle's point of view, this is referred to as a *passive transformation*.

Another way of explaining the distinction goes as follows. From the active point of view (situation 1), there is *one* observer looking at *two* events: the circle before and after the rotation. From the passive point of view (situation 2), there are *two* observers looking at *one* and the same event, but from a different perspective (read: coordinate system).

Both ways of expressing the symmetry of a circle are equivalent, but notice that the senses of rotation are opposite. If observers in the second mode tilt their head clockwise, the equivalent rotation in the first mode has to be counterclockwise (Figure 3.1), and vice versa. For this reason, it is tricky to mix the two ways of defining symmetry. Our choice will always be to leave the observers where they are—which means we keep the Cartesian frame fixed—and rotate the objects instead. In other words, we adhere to an *active interpretation* of symmetry transformations throughout this monograph.

At the risk of sounding overly repetitive, it bears emphasizing, once again, that during every (active) rotation of the circle, something is changing and something remains fixed. What is changing are the (x, y) coordinates of every point P on the circle. What remains the same is the image observed before and after the rotation. Note that a rotation leaves the distance between all points on the circle invariant, and hence also the radius of the circle.[3] Thinking *difference* and *identity* together is the essence of symmetry. The simple transformation of the position of the circle in space may look like a trivial children's game. The odd thing is that it is not. It involves the connection between matter and space, and carries us to the heart of physics, as we shall see.

In the next few sections, §3.2 through §3.6, we start by focusing on the changing (x, y) coordinates and express these transformations in mathematically precise terms. Then, in §3.7, we consider how the radius remains fixed (or invariant) during these rotations.

---

[2] This is also termed a *change of basis*.

[3] In Euclidean geometry, this is called an *isometry* (i.e., a transformation that moves points without changing the distances between them). Rotations are distinguished from other isometries by two additional properties: (1) they leave (at least) one point fixed (the origin) and (2) they leave *handedness* unchanged. In contrast, a *translation* moves every point: a *reflection* exchanges left- and right-hand ordering, and a *glide reflection* does both.

3.2 ROTATION OPERATORS

A rotation of the circle can be defined as an *operation* that moves all the points of the circle around the pivot point O over an angle ω. A positive value of ω corresponds to a counterclockwise rotation; a negative value, to a clockwise rotation, viewed from above the plane. Every rotation operation transforms the circle into a state that is indistinguishable from the original one; all rotations are, therefore, *symmetry operations* (or *covering operations*) of the circle.

We can represent each rotation operation by a *rotation operator*, denoted $R(\omega)$. The result of operating on any point P with rotation operator $R(\omega)$ is that the point P is moved to a new point P'. The former point can be considered to be the *target* of the operation whereas the latter is the *image*:

$$R(\omega)\,P = P'. \tag{3.1}$$

Definition 3.1 (Rotation operator): A *rotation operator*, denoted $R(\omega)$, is a mathematical entity that *operates* (acts) on a point P by rotating it over an angle ω about a fixed axis to the new point P'. ■

The rotation operator is thus defined as a mapping of points. Note that this operator affects all points in space. It acts *uniformly*. The origin is, of course, mapped onto itself. Furthermore, if P is lying on the circle, its image P' is also a point on the circle. Still, this is not sufficient to have an isometry of the circle. What distinguishes a symmetry operation from a random permutation of points is that the distances between points are not affected by the mapping.

The rotation angle ω is a real parameter, varying over the *periodic interval* $[0, 2\pi]$. This means that the angle starts over after it reaches 2π—the *modulus*. A rotation of 3π, for example, can be thought of as the sum of two successive rotations: one over 2π and another over 1π. Because a rotation of 2π corresponds to no rotation at all, the rotation of 3π is equivalent to one of π. One says the rotation angle ω is defined mod 2π (read: *modulo* 2π), where the mod operator ensures the rotation angle remains between 0 and 2π radians. Note that the number of possible rotations is infinite, as was mentioned in §1.2.3.

Simple as they may seem, these rotation operators, $R(\omega)$, have an unexpected power that they inherit from the special kind of mathematical structure to which they belong. They form the *elements* of a *symmetry group*.

3.3 THE AXIAL ROTATION GROUP

The set of all rotations in two-dimensional Euclidean space $\mathbb{R}^2$ forms an infinite continuous group—the *axial rotation group*. The elements of the group are the rotation operators $R(\omega)$, with $0 \leq \omega \leq 2\pi$. The product symbol $\star$ is defined to denote the consecutive application of two or more rotation operations. Having identified the nature of the elements and the product rule, it remains to be checked whether this particular set and operation satisfy all four group axioms:

1. *Closure*: For all $R(\omega_1)$ and $R(\omega_2)$ in $\mathcal{G}$, the result of their product is also an element of the group:

$$R(\omega_2) \star R(\omega_1) = R(\omega). \tag{3.2}$$

After all, $R(\omega_1)$ rotates the object over an angle ω_1, after which the object is rotated further over an angle ω_2 under the action of $R(\omega_2)$. This is equivalent to one rotation of $R(\omega)$ over an angle ω, where

$$\omega = \omega_1 + \omega_2 \pmod{2\pi}. \tag{3.3}$$

Note that this implies the group is Abelian, because $\omega_1 + \omega_2 = \omega_2 + \omega_1$—meaning, $R(\omega_2) \star R(\omega_1) = R(\omega_1) \star R(\omega_2)$.

2. *Associativity*: For all $R(\omega_1)$, $R(\omega_2)$, and $R(\omega_3)$ in $\mathcal{G}$, the following equation always holds:

$$[R(\omega_3) \star R(\omega_2)] \star R(\omega_1) = R(\omega_3) \star [R(\omega_2) \star R(\omega_1)]. \tag{3.4}$$

This can be explained by the fact that the total operation corresponds to a rotation over the sum of all three angles, which is, of course, associative:

$$[\omega_3 + \omega_2] + \omega_1 = \omega_3 + [\omega_2 + \omega_1]. \tag{3.5}$$

3. *Identity*: There exists an element E in $\mathcal{G}$ such that for every element $R(\omega)$ in $\mathcal{G}$,

$$E \star R(\omega) = R(\omega) \star E = R(\omega). \tag{3.6}$$

This identity element is simply the trivial rotation over 0°, $R(0)$, because

$$0 + \omega = \omega + 0 = \omega. \tag{3.7}$$

4. *Invertibility*: For each $R(\omega)$ in $\mathcal{G}$ there exists an element $R(\omega)^{-1}$ in $\mathcal{G}$ such that

$$R(\omega) \star R(\omega)^{-1} = R(\omega)^{-1} \star R(\omega) = E. \tag{3.8}$$

The inverse element corresponds to a rotation in the opposite sense; hence,

$$R(\omega)^{-1} = R(-\omega) \tag{3.9}$$

since

$$\omega + (-\omega) = -\omega + \omega = 0. \tag{3.10}$$

This verifies that the set of all two-dimensional rotation operators $R(\omega)$ forms a group—the axial rotation group.

3.4 TRANSFORMATIONS OF COORDINATES

So far, we have looked at the axial rotation group as an infinite set of physical operations denoted by the rotation operators $R(\omega)$. For each of these rotation operators $R(\omega)$, the corresponding mapping $P \mapsto P'$ may be expressed as a change of coordinates, where the coordinates of the target P are replaced by the coordinates of the image P'.

The aim of this section is to describe this *transformation of coordinates*. It will be convenient to express the transformation in matrix form, such that we can bring *matrix algebra* into play. This will become particularly useful in subsequent chapters when quantum mechanics enters the stage. A refresher of both vector and matrix algebra can be found in Appendices A and B.

Taking the origin O of the Cartesian system $\Sigma\left(x,y\right)$ as the fixed point, every point P can be addressed by a displacement vector $\mathbf{a}$ from the origin (Figure 3.2). The reason

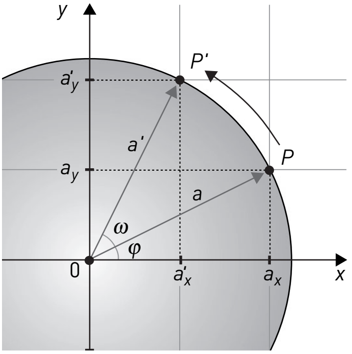


FIGURE 3.2 Counterclockwise rotation of point P through angle ω to the new point P' in a Cartesian coordinate system $\Sigma(x,y)$.

for working with *vectors* rather than with *points* comes from the fact that vectors can be represented by matrices. Suppose, for instance, that (a_x, a_y) are the coordinates of a random point P that represents the end point of a displacement vector **a** from the origin O. The vector **a** can then be written as a *column matrix*:

$$\mathbf{a} = \begin{bmatrix} a_x \\ a_y \end{bmatrix}. \tag{3.11}$$

Acting with the rotation operator $R(\omega)$ on P rotates this point counterclockwise over an angle ω to P' (i.e., $R(\omega)P = P'$). In vector notation, this corresponds to

$$R(\omega)\,\mathbf{a} = \mathbf{a}', \tag{3.12}$$

where $\mathbf{a}'$ is the position vector of P'. This can be rewritten in matrix notation as

$$R(\omega)\begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} a'_x \\ a'_y \end{bmatrix}. \tag{3.13}$$

The rotation operator $R(\omega)$ is said to induce a *linear transformation* of the coordinates. Now, to calculate the new coordinates (a'_x, a'_y) in terms of the old ones (a_x, a_y), the exact relationship between both sets of coordinates must be first identified; that is, one needs to understand the nature of the rotation operator $R(\omega)$ by expressing, in an analytical way, how it acts on the coordinates of P. This may be disclosed by reexpressing the coordinates as a function of the single parameter ϕ, which corresponds to the angle of elevation above the horizontal axis (Figure 3.2):

$$a_x = a\cos\phi; \tag{3.14}$$

$$a_y = a\sin\phi. \tag{3.15}$$

The angle ϕ is an angular coordinate with a range from 0 to 2π, and a is the magnitude of the position vector **a** (i.e., $a = |\mathbf{a}|$). When the point P with angular coordinate ϕ is rotated through the angle ω in a counterclockwise direction, it arrives at the new point P'. The elevation for the displaced point thus becomes

$$\phi' = \phi + \omega. \tag{3.16}$$

We may now use the parametric expressions in Eqs. (3.14–3.16) to find the a'_x and a'_y coordinates of the image point P'. All that is needed are the addition and subtraction formulas from basic trigonometry. One obtains

$$\begin{aligned} a'_x &= a\cos(\phi+\omega) \\ &= a\cos\phi\cos\omega - a\sin\phi\sin\omega \\ &= a_x\cos\omega - a_y\sin\omega. \end{aligned} \tag{3.17}$$

Analogously,

$$\begin{aligned} a'_y &= a\sin(\phi+\omega) \\ &= a\sin\phi\cos\omega + a\cos\phi\sin\omega \\ &= a_x\sin\omega + a_y\cos\omega. \end{aligned} \tag{3.18}$$

The results obtained in Eqs. (3.17) and (3.18) represent a linear transformation of the old coordinates (a_x, a_y) to the new coordinates (a'_x, a'_y). This may be conveniently expressed by means of a *rotation matrix*:

$$\mathbb{R}(\omega) = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}, \tag{3.19}$$

where all the coefficients of Eqs. (3.17) and (3.18) have been listed in a square array. To summarize, the rotation operator $R(\omega)$ turns any vector **a** into a new vector **a**′ that may be obtained via a matrix transformation of the original column vector:

$$R(\omega)\begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} a'_x \\ a'_y \end{bmatrix} = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}\begin{bmatrix} a_x \\ a_y \end{bmatrix}. \tag{3.20}$$

In vector notation,

$$R(\omega)\ \mathbf{a} = \mathbf{a}' = \mathbb{R}(\omega)\,\mathbf{a}. \tag{3.21}$$

Note that rotation matrices are always square. Although most applications involve rotations in two or three dimensions, rotation matrices can also be defined for higher order n-dimensional Euclidean spaces $\mathbb{R}^n$. In our discussion of the hydrogen atom, for example, use will be made of rotations in four dimensions, which can be represented by 4 × 4 rotation matrices (Chapter 9).

3.5 TRANSFORMATIONS OF COORDINATE FUNCTIONS

Now comes a subtle point! In the preceding section, we obtained an expression for the transformation matrix $\mathbb{R}(\omega)$, which acts on the coordinates of P and converts them into the coordinates of P' according to a linear transformation. Both sets of coordinates, (a_x, a_y) and (a'_x, a'_y), are ordinary *numbers* that refer to positions in space when a Cartesian coordinate system $\Sigma(x, y)$ is given. In this book, however, we are not concerned with numbers but with *functions.* A symbol such as $f(P)$ associates a value with point P, and the function f describes the distribution of these values over space.

A symmetry operation—say, R—that transforms the points can be claimed to induce a transformation of the function as well. Hence, we write

$$\hat{R}f(P) = f'(P). \tag{3.22}$$

Note that we have denoted the operation by a hat to distinguish it from operations acting on coordinates. That is, for each *coordinate operator* $R(\omega)$ there is a corresponding *function operator* $\hat{R}(\omega)$. There are several ways to define the connection between both. Let us opt again for the active view of symmetry, which implies that the “value travels with the, point." This means the value of the transformed function f' in the image point P' is equal to the value of the original function f in the original point P; hence,

$$f'\left(P'\right) = f(P). \tag{3.23}$$

Using Eq. (3.1), we obtain:

$$f'\,(RP) = f(P). \tag{3.24}$$

Now, to determine how a function changes, we want to compare the values of both functions *at the same point*. That is, we have to express the transformed function f' in terms of P, as in Eq. (3.22), rather than in terms of RP as we just did. Because this result is valid for any point P, we might as well apply it in point $R^{-1}P$. Let us therefore substitute P with $R^{-1}P$ in Eq. (3.24), yielding

$$f'\left(RR^{-1}P\right) = f(R^{-1}P). \tag{3.25}$$

And because $RR^{-1} = E$, we have

$$f'\,(P) = f(R^{-1}P). \tag{3.26}$$

Combining Eq. (3.22) with Eq. (3.26), we obtain the important relation

$$\hat{R}f\,(P) = f'\,(P) = f(R^{-1}P). \tag{3.27}$$

Definition 3.2 (Transformation of coordinate functions): The result of the action of an *operator* on a function f, denoted by the symbol $\hat{R}f$, is a new function f', which is obtained by applying the inverse coordinate transformation to the original function: $\hat{R}f\,(P) = f'\,(P) = f\left(R^{-1}P\right)$. ■

Example 3.1: Temperature distribution along a ring

The following example should clarify this definition. Consider a circular ring in the xy-plane of a Cartesian reference frame $\Sigma\left(x, y, z\right)$ (Figure 3.3). Now let us

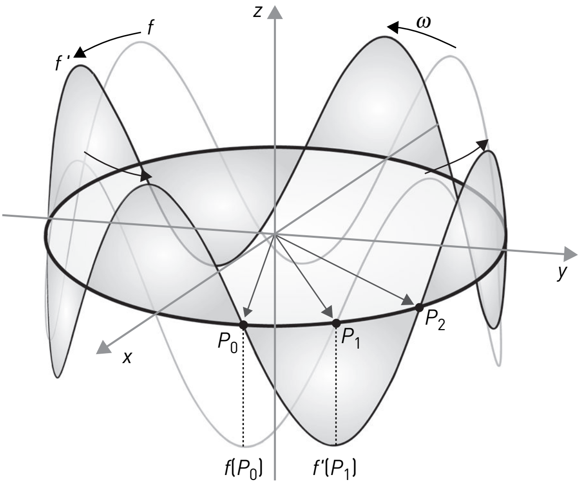


FIGURE 3.3 Temperature distribution along a circular ring before (f) and after (f') a counterclockwise rotation of the ring over angle ω.

associate a particular value with each point P on the ring—say, its temperature T. The temperature distribution along the ring is then given by a particular function $T = f(P)$. Let us assume, for instance, that as we walk along the ring, the temperature goes up and down in a periodic manner, between +20 °C and −20 °C, as shown in Figure 3.3 by the light-gray sinusoidal curve, where the z-axis denotes the temperature.

Now imagine rotating the ring around the z-axis in a counterclockwise fashion over an angle ω. Because the temperature "travels with the ring," the whole temperature distribution is rotated over the same angle ω in the same counterclockwise fashion. As seen from the fixed $\Sigma\left(x,y,z\right)$ reference frame, a new function $f'(P)$ is thus obtained, as shown by the dark-gray curve in Figure 3.3.

Now let us check the validity of Eq. (3.23), which says the value of the rotated function at the rotated point is equal to the value of the original function at the original point. Consider, for example, point P_1, for which the temperature was initially 0 °C, as indicated by the original function. This point was rotated to P_2, for which the new function indicates the temperature has indeed remained 0 °C, as required. That is,

$$f'(P_2) = \hat{R}(\omega) f\left(R(\omega)P_1\right) = f(P_1) = 0\,^\circ\mathrm{C}. \tag{3.28}$$

Alternatively, following Eq. (3.27), we could also consider the temperature of the *rotated* ring at P_1, which is −20 °C. This should equal the temperature of the original ring at P_0, which is the point obtained by rotating the original point P_1 clockwise over angle ω. In short,

$$f'(P_1) = \hat{R}(\omega) f(P_1) = f\left(R^{-1}(\omega)P_1\right) = f\left(R(-\omega)P_1\right) = f(P_0) = -20\,^\circ\mathrm{C}. \tag{3.29}$$

Example 3.2: Natural coordinate functions

As another example, let us look at a rather special kind of function known as *natural coordinate functions*. These simply correspond to the coordinate variables $\left(x,y\right)$. Clearly, these are functions. As an example, the value, which the function x associates with point P, is nothing else than the coordinate number a_x. Hence,

$$x(P) = a_x; \tag{3.30}$$

$$y(P) = a_y. \tag{3.31}$$

A convenient way to represent these functions is by means of two *osculating spheres*, as shown in Figure 3.4. This works as follows. Consider a given point P within a Cartesian reference frame $\Sigma\left(x,y\right)$ with x-coordinate $a_x = a\cos\phi$ and y-coordinate $a_y = a\sin\phi$ (Figure 3.4A). This point is represented in Figure 3.4C by the point P_x, which is lying on the radius vector to P but at a distance a_x from the origin (see also Figure 3.4A). For a point on the x-axis, P and P_x coincide, because in this case, $a_x = a$. On the other hand, by increasing the angle of elevation ϕ from 0 to $\pi/2$, keeping the radius a fixed, the P_x points describe a curved line, which ultimately approaches the origin when P reaches the y-axis, and hence $a_x = 0$ (Figure 3.4C). Similarly, when lowering the angle below the x-axis, we obtain a curve that approaches the origin from below for $\phi = -\pi/2$.

By a touch of magic, the curved lines in the interval $\phi \in [-\pi/2, +\pi/2]$ together form a perfect circle (in a perfect sphere when the z-coordinate joins in) centered on the positive x-axis at $x = a/2$, and with a radius equal to $a/2$ (Figure 3.4C).[4]

---

[4] As is seen in Figure 3.4C, the line segment for a given angle ϕ corresponds to a chord of the circle with radius $a/2$. The length of this chord is exactly $a_x = a\cos\phi$.

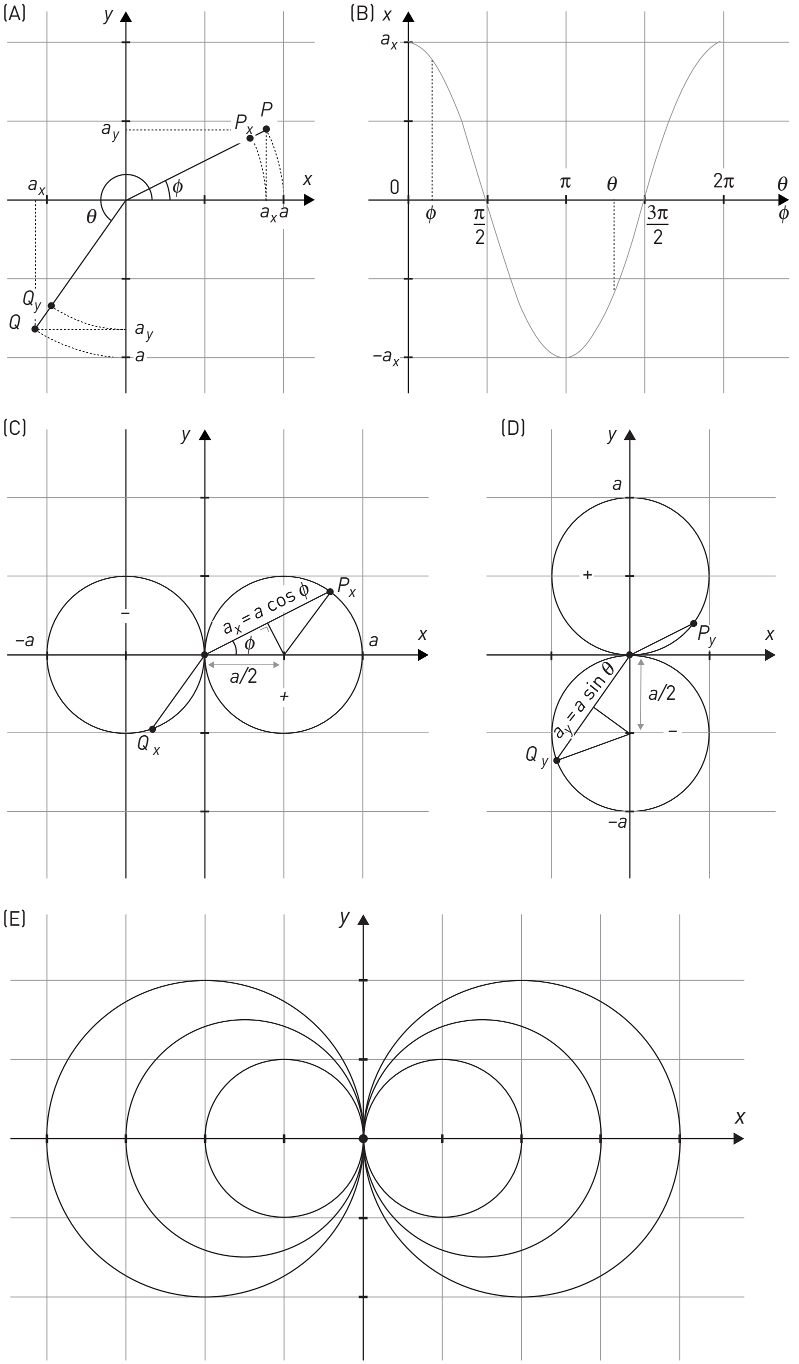


FIGURE 3.4 (A) Let *P* be a point with coordinates (a_x, a_y). (B) Here, we see the variation of a_x, with the angle ϕ for a fixed radius *a*. (C) All these a_x values are united in an angular plot. The positive a_x values are found in the interval $\phi \in [-\pi/2, +\pi/2]$ and describe a circle with radius $a/2$ centered at $x = a/2$. The negative values form a similar circle in the interval $\phi \in [+\pi/2, +3\pi/2]$, centered at $x = -a/2$. The two circles are tangent in the origin. (D) Here we see the analogous angular plot for the *y*-coordinate function. (E) This plot illustrates how the angular representation of the *x*-coordinate function varies for three different values of *a*.

The same circle is repeated along the negative x-axis, centered at $x = -a/2$, for $\phi \in [+\pi/2, +3\pi/2]$. The two spheres are tangent in the origin and form a kind of dumbbell structure, pointing along the x-direction. Chemists will recognize in this figure the common plot of the angular part of the atomic p_x-orbital. For the presentation of the whole function, osculating spheres should be drawn for each value of the radius, as is sketched in Figure 3.4E. The angular plot of the coordinate function for the y-coordinate is, of course, entirely analogous, but oriented along the y-axis (Figure 3.4D).

The plots make clear that the coordinate functions are oriented along the respective axes. We can think of them as polar vectors, with the head on the positive side of the dumbbell and the tail on the negative side. A counterclockwise rotation of the x-coordinate function over a quarter of a full angle turns the x-coordinate function into the y-coordinate function; the y-function, on the other hand, is turned into minus the x-function. These results are obtained immediately by simple visual inspection of the angular plots. Application of the general procedure for functional transformations reproduces these results in a rigorous way, as we will now show.

According to Eq. (3.23) we can write

$$\hat{R}(\omega)\, x\left(P'\right) = x(P) = a_x; \tag{3.32}$$

$$\hat{R}(\omega)\, y\left(P'\right) = y(P) = a_y. \tag{3.33}$$

Or, following Eq. (3.27),

$$\begin{aligned} x'(P) &= x(R(-\omega)P) \\ &= a\cos(\phi-\omega) \\ &= a\cos\phi\cos\omega + a\sin\phi\sin\omega \\ &= a_x\cos\omega + a_y\sin\omega \\ &= x(P)\cos\omega + y(P)\sin\omega. \end{aligned} \tag{3.34}$$

And similarly for $y'(P)$,

$$y'(P) = -x(P)\sin\omega + y(P)\cos\omega. \tag{3.35}$$

These results are valid for any point in space and thus fix a functional transformation, which means we can drop the point P altogether:

$$x' = x\cos\omega + y\sin\omega; \tag{3.36}$$

$$y' = -x\sin\omega + y\cos\omega. \tag{3.37}$$

This transformation, too, can be written as a matrix transformation. Because of the inverse relationship between the transformations of the coordinates and coordinate functions, the inverse matrix $\mathbb{R}(-\omega)$ is now required:

$$\hat{R}(\omega)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \mathbb{R}(-\omega)\begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.38}$$

Now, rotation matrices are special in that their inverse matrix simply corresponds to the transposed matrix (i.e., the matrix obtained by interchanging rows and columns). Hence, indicating transposition by a superscript T,

$$[\mathbb{R}(-\omega)]^{\mathrm{T}} = \mathbb{R}(\omega). \tag{3.39}$$

Taking the transpose of the full matrix equation means the coordinate functions are now arranged in a row vector, which precedes the transposed transformation matrix:

$$\hat{R}(\omega)\begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}\mathbb{R}(\omega). \tag{3.40}$$

The natural coordinate functions form what is called a *function space*. The choice to write the components of a function space in a row vector notation has two advantages: (1) the same rotation matrix $\mathbb{R}(\omega)$ can be used for coordinates and coordinate functions, notwithstanding their opposite transformations; and (2) the mapping of the set of rotation matrices on the rotation group is an isomorphism, as we will see in the next paragraphs.

3.6 MATRIX REPRESENTATIONS

The question arises regarding what is the connection between the group of rotation operators and the set of rotation matrices? In §3.6.1, we first consider the relation with the coordinate operators $R(\omega)$. In §3.6.2, we then repeat the same consideration for the function operators $\hat{R}(\omega)$. Although the treatment in this section applies to rotation operators, it is, in fact, more general and applies to any set of symmetry operations. In the following, symmetry operations are therefore denoted more generally by the symbols R, S, T, and so on.

3.6.1 Matrix representation of coordinate operators *R*

Suppose one applies two operators, S and R, consecutively to a point P. This is expressed by a right-justified equation, which means the operator on the right acts first:

$$(S \star R)\, P = S \star (R\, P) = S\, P' = P''. \tag{3.41}$$

This is illustrated in Figure 3.5. The corresponding coordinate transformation may be obtained by combining the two separate transformations:

$$R\begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} a'_x \\ a'_y \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}\begin{bmatrix} a_x \\ a_y \end{bmatrix} = \mathbb{R}\begin{bmatrix} a_x \\ a_y \end{bmatrix}; \tag{3.42}$$

$$S\begin{bmatrix} a'_x \\ a'_y \end{bmatrix} = \begin{bmatrix} a''_x \\ a''_y \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}\begin{bmatrix} a'_x \\ a'_y \end{bmatrix} = \mathbb{S}\begin{bmatrix} a'_x \\ a'_y \end{bmatrix}. \tag{3.43}$$

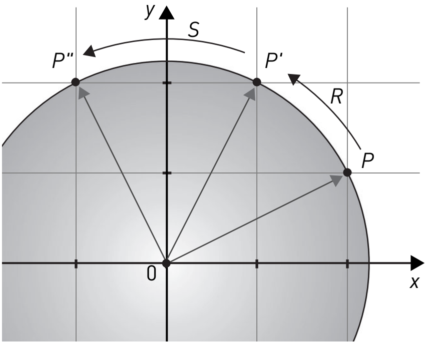


FIGURE 3.5 When the coordinate operator $S \star R$ acts on a point P, P is first rotated to P' under the action of R and P' is then further rotated to P'' by S.

Here, it behooves us to remember that the symmetry operations concern uniform transformations. Hence, the matrices are independent of the coordinates of the points on which they act. The matrix $\mathbb{S}$ can thus be used to rotate point P' to P'' further. As an example, the a''_x number is calculated from a'_x and a'_y, which in turn are transformations of a_x and a_y:

$$\begin{aligned} a''_x &= s_{11}a'_x + s_{12}a'_y \\ &= s_{11}\left(r_{11}a_x + r_{12}a_y\right) + s_{12}\left(r_{21}a_x + r_{22}a_y\right) \\ &= \left(s_{11}r_{11} + s_{12}r_{21}\right)a_x + \left(s_{11}r_{12} + s_{12}r_{22}\right)a_y. \end{aligned} \tag{3.44}$$

Here, s_{ij} and r_{kl} denote matrix elements of the respective matrices. As we can see, the transformation coefficients are obtained by performing the matrix product of the individual matrices R and S (as defined in Appendix B.4):

$$\begin{bmatrix} a''_x \\ a''_y \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \mathbb{S} \times \mathbb{R} \begin{bmatrix} a_x \\ a_y \end{bmatrix}. \tag{3.45}$$

Because the operators are part of a group, their product $S \star R$ is another symmetry operation, which we can represent as T. This operation maps P onto P'' directly; that is, $TP = P''$ or, in matrix form,

$$\begin{bmatrix} a''_x \\ a''_y \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \mathbb{T} \begin{bmatrix} a_x \\ a_y \end{bmatrix}. \tag{3.46}$$

Combining these results leads to the following mapping:

$$T = S \star R \quad \longleftrightarrow \quad \mathbb{T} = \mathbb{S} \times \mathbb{R}. \tag{3.47}$$

This result indicates that the set of matrices associated with a set of symmetry operators forms a representation of the symmetry group. The combination rule is represented by the matrix product. The order of the operations is also preserved, which implies that the mapping is an isomorphism (or homomorphism). This mapping is called a *matrix representation* of the group. The consecutive action of the physical operators in real space is replaced by matrix multiplications. The unit operation and the reverse rotation find their logical mathematical counterparts in the unit and inverse matrices.

3.6.2 Matrix representation of function operators $\hat{R}$

We now repeat the derivation for the function operators. There are two important changes: (1) the coordinate functions are represented in rows, and not in columns, and (2) the way the operators act on the components of the function space is different. Linear operators in a function space leave the numerical coefficients preceding the functions unchanged and act directly on the basic functional components following the coefficients.

Consider, as an example, the action of the operator product $\hat{S} \star \hat{R}$ on the function space of natural coordinate functions:

$$\left(\hat{S} \star \hat{R}\right) \begin{bmatrix} x & y \end{bmatrix} = \hat{S} \star \left(\hat{R} \begin{bmatrix} x & y \end{bmatrix}\right) = \hat{S} \begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x'' & y'' \end{bmatrix}. \tag{3.48}$$

The action of both operators is given by

$$\hat{R} \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{R}; \tag{3.49}$$

$$\hat{S}\begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}\mathbb{S}. \tag{3.50}$$

The action of the $\hat{S}$ operator on the function x' in Eq. (3.48) is then given by

$$\begin{aligned} \hat{S}x' &= \hat{S}\left(r_{11}x + r_{21}y\right) \\ &= r_{11}\hat{S}x + r_{21}\hat{S}y \\ &= r_{11}\left(s_{11}x + s_{21}y\right) + r_{21}\left(s_{12}x + s_{22}y\right) \\ &= \left(s_{11}r_{11} + s_{12}r_{21}\right)x + \left(s_{21}r_{11} + s_{22}r_{21}\right)y. \end{aligned} \tag{3.51}$$

The resulting matrix transformation is the same product matrix obtained previously—in matrix notation:

$$\begin{aligned} \hat{S}\star\hat{R}\begin{bmatrix} x & y \end{bmatrix} &= \hat{S}\begin{bmatrix} x & y \end{bmatrix}\mathbb{R} \\ &= \begin{bmatrix} x & y \end{bmatrix}\mathbb{S}\times\mathbb{R} \\ &= \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}. \end{aligned} \tag{3.52}$$

Denoting the operator product $\hat{S}\star\hat{R}$ by the single operator $\hat{T}$ yields

$$\hat{T}\begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}\mathbb{T} = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}. \tag{3.53}$$

Hence, symmetry transformations in function space lead again to a mapping of the group that preserves the order of the operations:

$$\hat{T} = \hat{S}\star\hat{R} \quad \longrightarrow \quad \mathbb{T} = \mathbb{S}\times\mathbb{R}. \tag{3.54}$$

If the components of the function space were ordered in a column vector, the order would have been reversed and we would have obtained an *anti-isomorphism*.

3.7 THE ORTHOGONAL GROUP O(2)

One of the principal recurring themes of this book is that symmetries refer to some property that remains invariant. Indeed, instead of deriving the symmetry group from the physical covering operations that map a circle onto itself, as we did in the previous sections, we could also start from the invariance principle, which requires the length of the radius vector to remain unchanged. As we will see in this section, this approach provides an alternative definition of the symmetry group as a set of *orthogonal transformations*.

3.7.1 Symmetry and invariance

Eqs. (3.20) and (3.21) describe the exact relationship between the initial $\left(a_x, a_y\right)$ and the final $\left(a'_x, a'_y\right)$ coordinates of a point P being rotated over an angle ω around the origin O. Now, despite the fact that P changes position, the distance OP never alters during these rotations. That is to say, when the position vector $\mathbf{a}$ is turned into $\mathbf{a}'$, its direction changes, but its magnitude remains invariant:

$$|\mathbf{a}| = |\mathbf{a}'|. \tag{3.55}$$

This conservation of distance is called an *invariance property*. Let us rewrite the previous equation in matrix form. According to the Pythagorean theorem, $a_x^2 + a_y^2 = \mathbf{a}^2$; therefore, the magnitude of **a** equals

$$|\mathbf{a}| = \sqrt{a_x^2 + a_y^2}. \tag{3.56}$$

Since $a_x^2 + a_y^2$ is also the dot product of **a** with itself (see Appendix A), this equation may be rewritten as follows:

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}. \tag{3.57}$$

Restating the invariance property in Eq. (3.55) in terms of Eq. (3.57) and squaring everything yields:

$$\mathbf{a} \cdot \mathbf{a} = \mathbf{a}' \cdot \mathbf{a}'. \tag{3.58}$$

Following Eq. (B.15), these inner products may be reformulated in matrix form:

$$\mathbf{a}^{\mathrm{T}}\mathbf{a} = \mathbf{a}'^{\mathrm{T}}\mathbf{a}', \tag{3.59}$$

where **a** and **a**′ represent two column vectors as defined in Eq. (3.11). For each of these column vectors, there is a corresponding *dual vector*, $\mathbf{a}^{\mathrm{T}}$ or $\mathbf{a}'^{\mathrm{T}}$, that is defined as the transposed form or row form of **a** and **a**′ (cf. Eq. (B.14)):

$$\mathbf{a}^{\mathrm{T}} = \begin{bmatrix} a_x & a_y \end{bmatrix}; \quad \mathbf{a}'^{\mathrm{T}} = \begin{bmatrix} a'_x & a'_y \end{bmatrix}. \tag{3.60}$$

Now let us introduce a transformation, A, represented by the matrix $\mathbb{A}$, that alters the radius vectors but keeps the distance the same. The **a** and $\mathbf{a}^{\mathrm{T}}$ transform as

$$A\mathbf{a} = \mathbf{a}' = \mathbb{A}\mathbf{a}; \tag{3.61}$$

$$A\mathbf{a}^{\mathrm{T}} = \mathbf{a}'^{\mathrm{T}} = [\mathbb{A}\mathbf{a}]^{\mathrm{T}} = \mathbf{a}^{\mathrm{T}}\mathbb{A}^{\mathrm{T}}. \tag{3.62}$$

The dual vector $\mathbf{a}^{\mathrm{T}}$ is thus rotated by the transposed matrix $\mathbb{A}^{\mathrm{T}}$. Note that the order of $\mathbf{a}^{\mathrm{T}}$ and $\mathbb{A}^{\mathrm{T}}$ has been reversed in Eq. (3.62) as compared to Eq. (3.61), in accordance with the familiar rule for the transpose of a matrix product as described in Eq. (B.16). Based on Eqs. (3.61) and (3.62), the invariance relation in Eq. (3.59) now yields

$$\mathbf{a}^{\mathrm{T}}\mathbf{a} = \mathbf{a}^{\mathrm{T}}\ \mathbb{A}^{\mathrm{T}}\mathbb{A}\ \mathbf{a}. \tag{3.63}$$

As we saw before, invariance should refer not only to the length of the radius vector, but should also imply that the distances between all the points of the object remain unchanged. Hence, a more general invariance principle reads

$$\mathbf{a}_1^{\mathrm{T}}\mathbf{a}_2 = \mathbf{a}_1^{\mathrm{T}}\ \mathbb{A}^{\mathrm{T}}\mathbb{A}\ \mathbf{a}_2. \tag{3.64}$$

Here, vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are directed to different points on the circle. For this general set of equations to hold, $\mathbb{A}^{\mathrm{T}}$ must correspond to the inverse of $\mathbb{A}$:

$$\mathbb{A}^{\mathrm{T}}\mathbb{A} = \mathbb{I} = \mathbb{A}\mathbb{A}^{\mathrm{T}}, \tag{3.65}$$

where $\mathbb{I}$ is the identity matrix. Matrices that have this property are called *orthogonal matrices* (see Appendix B). Thus, the set of orthogonal matrices qualifies as transformations that obey the invariance principle. The appearance of matrix orthogonality as a unique and sufficient criterion for invariance raises several intriguing questions. First of all, are *all* orthogonal matrices, by definition, rotation matrices (see §3.7.2)? Second, do orthogonal matrices form a group under matrix multiplication (see §3.7.3)? And, if so, how does the structure of this group relate with the axial rotation group (see §3.7.3)? All these questions can be solved algebraically, as we will show in the following paragraphs.

3.7.2 Proper and improper rotation matrices

In the previous section, the conclusion was reached that all rotation matrices must obey the orthogonality condition. But what about the opposite? Are *all* orthogonal matrices also rotation matrices? A straightforward way to find this out is by constructing a system of equations that expresses the orthogonality of a 2×2 matrix $\mathbb{A}$ with four real unknowns: a, b, c, and d. According to Appendix B, orthogonality means

$$\begin{aligned} \mathbb{A}\mathbb{A}^{\mathrm{T}} = \mathbb{I} \Rightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \\ \mathbb{A}^{\mathrm{T}}\mathbb{A} = \mathbb{I} \Rightarrow \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{3.66}$$

These two matrix products are equivalent to a set of six equations with four unknowns:

$$\begin{array}{llll} i. & a^2 + b^2 = 1; & iv. & b^2 + d^2 = 1; \\ ii. & c^2 + d^2 = 1; & v. & ac + bd = 0; \\ iii. & a^2 + c^2 = 1; & vi. & ab + cd = 0. \end{array} \tag{3.67}$$

A general solution for Eq. (3.67.*iii*) is given by

$$a = \cos\omega; \quad c = \sin\omega, \tag{3.68}$$

since $\cos^2\omega + \sin^2\omega = 1$. Eqs. (3.67.*i*)—(3.67.*iii*) yields $b^2 = c^2$; Eqs. (3.67.*i*)—(3.67.*iv*) gives $a^2 = d^2$. It thus follows that

$$b = \pm\sin\omega; \quad d = \pm\cos\omega. \tag{3.69}$$

Combination of this result with the final two equations in Eq. (3.67) demonstrates that the values of b and d are interdependent. That is, either

$$b = -\sin\omega, \quad d = +\cos\omega; \tag{3.70}$$

or

$$b = +\sin\omega, \quad d = -\cos\omega. \tag{3.71}$$

There are, consequently, two general forms of orthogonal matrices:

$$\mathbb{A} = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}; \tag{3.72}$$

$$\mathbb{B} = \begin{bmatrix} \cos\omega & \sin\omega \\ \sin\omega & -\cos\omega \end{bmatrix}. \tag{3.73}$$

Although both types are *unimodular matrices* (i.e., square matrices with a determinant of $+1$ or -1), the sign of their determinant differs:

$$\begin{aligned} \det\mathbb{A} &= ad - bc = cos^2\omega + sin^2\omega = 1; \\ \det\mathbb{B} &= ad - bc = -(cos^2\omega + sin^2\omega) = -1. \end{aligned} \tag{3.74}$$

When comparing the matrices of type $\mathbb{A}$ with the rotation matrices $\mathbb{R}(\omega)$ of the previous section, they are seen to correspond (compare Eqs. (3.19) and (3.72)). A second algebraic property of rotation matrices has thus been disclosed. That is, to qualify as a rotation matrix $\mathbb{R}(\omega)$, any matrix $\mathbb{A}$ should be both orthogonal ($\mathbb{A}^{\mathrm{T}} = \mathbb{A}^{-1}$) *and* have a determinant equal to 1 ($\det\mathbb{A} = 1$). All matrices satisfying these two conditions act as rotation matrices, and they are therefore termed *proper* rotation matrices.

But what about the other type of orthogonal matrices in Eq. (3.73) with a determinant equal to -1? Although all physical rotations are expressed by means of the proper rotation matrices $\mathbb{A}$, matrices of type $\mathbb{B}$ are seen to describe physical *reflections*, and they are therefore called *improper* rotation matrices. The simplest possible case is obtained for $\omega = 0$:

$$\sigma_h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \quad \det \sigma_h = -1. \tag{3.75}$$

The matrix σ_h corresponds to a mirror operation across the horizontal x-axis, keeping the x-coordinate fixed, but inverting the y-coordinate:

$$\sigma_h \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} a'_x \\ a'_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} a_x \\ -a_y \end{bmatrix}. \tag{3.76}$$

In the following, we will not make use of reflections, except for this horizontal mirror operation.

3.7.3 Orthogonal groups: O(2) and SO(2)

The orthogonal group O(2)

The set of all 2 × 2 orthogonal matrices (consisting of all proper and improper rotation matrices) is said to form an *orthogonal group* under the operation of matrix multiplication; this group is denoted by the symbol O(2), with “O” for Orthogonality and “2” because we are dealing with 2 × 2 matrices. Let us verify this assertion by checking the four group postulates:

1. *Closure*: For all $\mathbb{U}$ and $\mathbb{V}$ in O(2), the result of the operation $\mathbb{UV}$ is also in O(2). The proof consists of showing that $\mathbb{UV}$ is a matrix that obeys the orthogonality condition. Since $\mathbb{U}$ and $\mathbb{V}$ are both orthogonal matrices (i.e., $\mathbb{U}\mathbb{U}^{\mathrm{T}} = \mathbb{U}^{\mathrm{T}}\mathbb{U} = \mathbb{I}$ and $\mathbb{V}\mathbb{V}^{\mathrm{T}} = \mathbb{V}^{\mathrm{T}}\mathbb{V} = \mathbb{I}$):

$$\begin{aligned} (\mathbb{UV})(\mathbb{UV})^{\mathrm{T}} &= \mathbb{UV}\mathbb{V}^{\mathrm{T}}\mathbb{U}^{\mathrm{T}} = \mathbb{U}\mathbb{I}\mathbb{U}^{\mathrm{T}} = \mathbb{U}\mathbb{U}^{\mathrm{T}} = \mathbb{I}; \\ (\mathbb{UV})^{\mathrm{T}}(\mathbb{UV}) &= \mathbb{V}^{\mathrm{T}}\mathbb{U}^{\mathrm{T}}\mathbb{UV} = \mathbb{V}^{\mathrm{T}}\mathbb{I}\mathbb{V} = \mathbb{V}^{\mathrm{T}}\mathbb{V} = \mathbb{I}, \end{aligned} \tag{3.77}$$

   where use was made of the transposition of a matrix product (see Eq. (B.16)).
2. *Associativity*: For all $\mathbb{U}$, $\mathbb{V}$, and $\mathbb{W}$ in O(2), the equation $(\mathbb{UV})\mathbb{W} = \mathbb{UVW} = \mathbb{U}(\mathbb{VW})$ holds because matrix multiplication is associative.
3. *Identity element*: There exists an element $\mathbb{E}$ in O(2) such that for every element $\mathbb{U}$ in O(2), the equation $\mathbb{EU} = \mathbb{UE} = \mathbb{U}$ holds. This identity element $\mathbb{E}$ is simply the 2 × 2 unit matrix $\mathbb{I}$.
4. *Inverse element*: For each $\mathbb{U}$ in O(2), there exists an element $\mathbb{U}^{-1}$ in O(2) such that $\mathbb{U}\mathbb{U}^{-1} = \mathbb{U}^{-1}\mathbb{U} = \mathbb{I}$. This last requirement is satisfied because all square matrices with nonzero determinants have an inverse matrix. For orthogonal matrices, this inverse is equal to the transposed matrix (i.e., $\mathbb{U}^{-1} = \mathbb{U}^{\mathrm{T}}$), and since $\mathbb{U}^{\mathrm{T}}$ is orthogonal, $\mathbb{U}^{-1} \in \mathrm{O}(2)$.

The special orthogonal group SO(2)

If we restrict the set of orthogonal matrices in O(2) to the subset of proper rotation matrices with determinant +1, it is easy to verify that this subset forms a group as well. Although the proof proceeds along the same lines as those just presented, the argument also relies on the additional property that the determinant of a matrix product is equal to the product of the determinants:

$$\det \mathbb{U}\mathbb{V} = \det \mathbb{U} \det \mathbb{V}. \tag{3.78}$$

The importance of this equation follows from the fact that the set of rotation matrices should satisfy the axiom of closure to qualify as a group. That is, the product of two rotation matrices $\mathbb{U}$ and $\mathbb{V}$ (both with unit determinant) should result in a new rotation matrix $\mathbb{U}\mathbb{V}$ (again with unit determinant). This is verified in Eq. (3.78).

The subset of proper rotation matrices thus forms a group—the *special orthogonal group* in two dimensions, denoted SO(2). In other words, the axial rotation group is *isomorphic* to the special orthogonal group SO(2). The group is referred to as *special* because the set of orthogonal matrices has been restricted to those with unit determinant.

Following Definition 2.3 in §2.5, the group SO(2) is said to form a *subgroup* of O(2). This relation can be represented by the following *symmetry breaking chain*:

$$\mathrm{O}(2) \supset \mathrm{SO}(2). \tag{3.79}$$

Notice also that the subgroup SO(2) is Abelian (as was shown in §3.3), whereas the parent group O(2) is not.

The subset of improper rotation matrices (i.e., reflection matrices with determinant −1), on the other hand, does *not* form a subgroup of O(2) because the product of two reflections is not a reflection, but a rotation—thus violating the closure property of groups. For instance, the matrix corresponding to a mirror operation across the vertical y-axis is

$$\sigma_v = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \det \sigma_v = -1. \tag{3.80}$$

The product of Eqs. (3.80) and (3.75) is then

$$\begin{gathered} \sigma_v \sigma_h = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \mathbb{R}(\pi); \\ \det(\sigma_v \sigma_h) = 1, \end{gathered} \tag{3.81}$$

where $\mathbb{R}(\pi)$ stands for a rotation over 180°. In terms of determinants, $(-1)(-1) = +1$.

The entire subset of improper rotations is generated by multiplying all the elements of the SO(2) group with a single mirror operation—say, σ_h. We denote the resulting set as $\sigma_h\mathrm{SO}(2)$. We may thus write the group O(2) of rotations and reflections as the following sum:

$$\mathrm{O}(2) = \mathrm{SO}(2) + \sigma_h \mathrm{SO}(2). \tag{3.82}$$

The subgroup SO(2) is at the basis of a *partitioning* of O(2) into two halves (Figure 3.6): the subgroup itself and the complementary set of reflections. The elements of the partition of a group over a subgroup are called *cosets*.

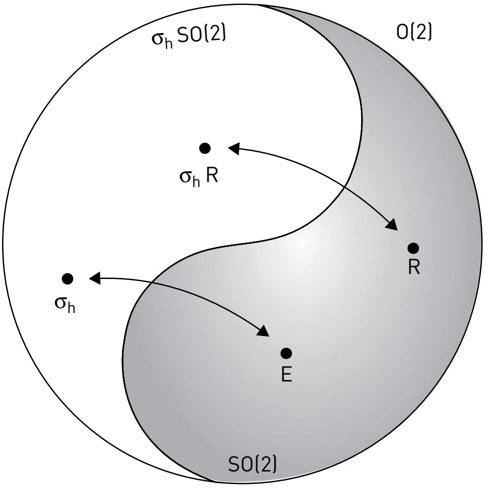


FIGURE 3.6 Schematic representation of the coset decomposition of the orthogonal group O(2). The group O(2), represented here by a full circle, can be thought of as the union of the SO(2) subgroup of rotations (gray region) and the coset σ_h SO(2) of reflections (white region). To each element from SO(2) then corresponds an element from the coset σ_h SO(2).

Definition 3.3 (Coset): A *coset* $g\mathcal{H}$ of a subgroup $\mathcal{H}$ of $\mathcal{G}$ is the set of all elements obtained by multiplying every element of $\mathcal{H}$ by a given element g of $\mathcal{G}$:

$$g\mathcal{H} = \left\{ g \star h_i \mid g \in \mathcal{G}, h_i \in \mathcal{H} \right\}. \tag{3.83}$$

For $g \in \mathcal{H}$, the coset coincides with $\mathcal{H}$. This is the *trivial coset.* ▪

In the case of the O(2) group, only one nontrivial coset is found. In other words, every reflection operation $g \notin \mathrm{SO}(2)$ will give rise to the same coset $\sigma_h \mathrm{SO}(2)$.

4 The SO(2) group

I would like to compare {Sophus Lie} to a Scout in the primeval forest who, when others had given up trying to break through the undergrowth in desperation, always knew how to find a way, the way that affords the best views of unfamiliar, yet romantic mountains and valleys.
–Friedrich Engel (November 14, 1899)

Norwegian mathematician Sophus Lie (1842–1899)—a fellow countryman of Niels Hendrik Abel—first met Felix Klein (1849–1925) in winter 1869 during a visit to Berlin (Figure 4.1).[1] At that time, Berlin was considered the center of the mathematical world, with the towering figures of Ernst Kummer, Leopold Kronecker, and Karl Weierstrass dominating the mathematical scene. Lie and Klein were only at the verge of their mathematical career by then—Lie was twenty-seven and Klein was seven years his junior—but the two men got along immediately and became lifelong friends.

In spring 1870, Lie and Klein traveled to Paris, where they first met Jean-Gaston Darboux (1842–1917) and Camille Jordan (1838–1922), whose book on Galois' group theory had just been published. Both Lie and Klein had toyed with the group concept before on a rather intuitive basis, but it was Jordan who really opened their eyes (and minds) to the central importance of group theory. All of this came to a sudden end, however, on July 19, 1870, when the Franco-Prussian War broke loose. Klein,

---

[1] The historical prelude to this chapter is based on the truly fascinating and definitive biography of Sophus Lie, written by Arild Stubhaug. All the quotes are also taken from his book. See A. Stubhaug. *The Mathematician Sophus Lie: It Was the Audacity of My Thinking*. Berlin: Springer-Verlag, 2002. A shorter account of Sophus Lie's life and work can be found in B. Fritzsche. "Sophus Lie: A Sketch of His Life and Work." *Journal of Lie Theory* 9 (1999), pp. 1–38.

FIGURE 4.1 Norwegian mathematician Sophus Lie (1842–1899) in an oil painting by Erik Theodor Werenskiold, from the year 1902.

a Prussian, felt forced to return to Berlin, where he joined the military service. One month later, Lie began to feel the heat around the corner as well, and he left the City of Lights too.

Equipped with nothing more than a backpack, Lie set out on a long hiking trip to Italy,[2] where he hoped to meet some of the great Italian mathematicians, such as Luigi Cremona (1830–1903) in Milan. But, having hiked a mere 60 km, from Paris to Fontainebleau, Lie got arrested by the French on suspicion of being a German spy. His backpack was searched by the gendarmes, who mistook his mathematical notes for top secret, cryptically encoded messages. The event made Lie instantly famous across Norway. The headlines of the local newspapers read: “Norwegian Scientist Jailed as German Spy.”[3] But Lie, the ever-strong blond and fearless Viking, didn’t seem too distressed about the whole situation. One day, Lie casually asked one of the guards what usually happened to German prisoners. “We usually shoot them at six in the morning!” replied the guard somewhat ironically.

After one month of imprisonment, Lie was finally released thanks to the intervention of Darboux, who convinced the French soldiers of the innocence of Lie’s mathematical notes. Lie took a train via Switzerland to Italy and returned to Norway. On his journey back, Lie passed through Germany, where he was welcomed once again by Klein in Berlin.

The two men thrived in each other’s company and spent hours discussing mathematics. During these exciting talks, Lie and Klein stumbled upon the idea of studying the properties of a mathematical object by looking at its *symmetry group*. The many fruitful conversations they had shared with Jordan, back in Paris, had ripened in their minds and they were both convinced of pursuing the group theoretical

---

[2] Lie loved to hike—in particular, in his “beloved mountains” in Norway. One weekend, Lie decided to visit his father in Moss because he needed a book from home. Legend says he hiked nearly 60 km, from Christiania to Moss, to find out that his father was not at home. Shrugging his shoulders, Lie turned his back and happily hiked all the way back. Mathematical folklore even says that whenever it started to rain when Lie was hiking, he simply took off his clothes and stuffed them in his backpack.

[3] Stubhaug, *The Mathematician Sophus Lie*, p. 14.

path. They decided that Lie would study "*infinitesimal groups*" (now called *infinite continuous groups*), whereas Klein would focus on discrete groups.

For Klein, this ultimately resulted in his "Erlanger Programm," whereas for Lie, the study of infinite continuous groups culminated in the publication of his three-volume magnum opus titled *Theorie der Transformationsgruppen*, which he had written with the help of Friedrich Engel (1861–1941) during the period 1884 to 1893. In a letter to a colleague, Lie once wrote: "I am certain, absolutely certain, that, at some point in the future, these theories will be recognized as fundamental."[4] His prophecy was soon to be fulfilled.

Lie's work had opened up a new branch of mathematics that was about to revolutionize the mathematical landscape, just as Galois had done half a century before. Lie's transformation groups became known as *Lie groups*, after Hermann Weyl (1885–1955) first coined the term during one of his seminars in Princeton during the mid 1930s. "Little by little," said Jean Dieudonné (1906–1992), a French mathematician and historian of mathematics, "it became obvious that the most unexpected theories, from arithmetic to quantum physics, came to encircle [Lie's theory] like a gigantic axis."[5]

Chapter outline

The aim of this fourth chapter is to take you on a whistle-stop tour through the "romantic mountains and valleys" of modern *Lie group theory* to which Engel referred in the opening quote of this chapter. Most of the concepts, which we will examine in this chapter, will be worked out in greater detail in the following chapters.

Lie groups are a special kind of *infinite continuous group* whose elements can be parametrized smoothly and analytically.[6] Let us therefore initiate this chapter with a description of infinite continuous groups in §4.1, before specializing to the central concept of a Lie group in §4.2.

In the following section, §4.3, the *infinitesimal generators* of a Lie group are introduced. The usefulness of this powerful concept comes from the fact that all elements from a particular Lie group can be *generated* from this small set of generators. As a consequence, we no longer have to deal with the infinite set of group elements; we can focus all our attention on a study of the finite set of generators that give rise to the Lie group and therefore embody most of the structure of Lie groups.

However, although all elements from the Lie group can be obtained from the generators by their repeated application, the generators themselves do not belong to the Lie group. They instead form a basis for a different kind of algebraic structure that is commonly known as a *Lie algebra*. A closer study of the deep connection between Lie algebras and their corresponding Lie groups will be undertaken in the next chapter.

Most of the concepts in this chapter are illustrated with the help of the two-dimensional axial rotation group. This group is isomorphic to the group of special orthogonal 2 × 2 matrices, denoted SO(2), and is the simplest known one-parameter Lie group. Therefore, it provides the perfect starting point for our discussion of

---

[4] Stubhaug, *The Mathematician Sophus Lie*, p. 18.

[5] Quoted from Stubhaug, *The Mathematician Sophus Lie*, p. 115.

[6] Don't worry about this technical language. We will examine each concept, one at a time, and with sufficient care to make them crystal clear.

Lie groups and infinitesimal generators. As we shall see in §4.3, the SO(2) group has only one infinitesimal generator. Consequently, its Lie algebra lacks the rich structure we will encounter later in problems with higher symmetry. Nevertheless, the single infinitesimal operator of SO(2) will be seen to play the protagonist part when quantum mechanics enters the stage in § 4.4.

In the final section of this chapter, §4.5, the simple model of an electron orbiting on a ring is presented, which embodies SO(2) symmetry and provides a model description for the electronic structure of aromatic molecules. These molecules can be classified by a simple sequence of whole numbers—2, 6, 10, 14, 18, and so on—known as the *Hückel numbers.* These numbers refer directly to the Lie structure of the SO(2) group. In this way, we get a first glimpse of how a symmetry group can be realized in a sequence of compounds. Farther up the road, more intricate and intriguing number sequences are waiting for similar encounters.

4.1 INFINITE CONTINUOUS GROUPS

Recall that the notions of finite and infinite order groups were introduced in Chapter 2. In addition, a further distinction was made between infinite discrete and infinite continuous groups. Their definitions are given in §2.1.

4.1.1 The nature of infinite continuous groups

Although classical group theory is mainly concerned with finite and infinite discrete groups, the theory of Lie groups, on the other hand, deals with a special kind of infinite continuous groups, about which we will have more to say in §4.2. When dealing with finite and infinite discrete groups, abstract group theory (or *pure* group theory, if you like) provides a sufficiently broad framework such that no other parts of mathematics are usually needed. Infinite continuous groups, on the other hand, are richer in nature; they exhibit a number of additional properties that provide a point of contact with other branches of mathematics, such as ordinary and partial differential algebra, topology, geometry, and so forth. Naturally, this renders the subject matter mathematically more demanding. Considerable care is, therefore, required when formulating the general framework of continuous groups. The goal, however, is to stick to the theoretical minimum, and leave most topological and geometric excursions aside.

The principal reason for so doing comes from the fact that all continuous groups we will encounter throughout this book are isomorphic to (read: can be represented as) *matrix groups* whose additional topological and geometrical properties are well known and rather straightforward. These groups are commonly named *linear* or *classical* Lie groups, because the elements of these groups ($n \times n$ matrices) represent *linear transformations* of coordinates or coordinate functions in n-dimensional space. Most Lie groups that are of interest to physicists and chemists are linear.

4.1.2 Parameters of continuous groups

As mentioned in the definition of infinite continuous groups (§2.1), the elements $\hat{R}$ of a continuous group $(\mathcal{G}, \star)$ are labeled by a finite set of n independent *parameters* r_i ($i = 1, \ldots, n$), all of which vary continuously over a particular interval. The parameters are always chosen to form a set that is both *necessary* and *sufficient*. That is to say, the particular choice of n parameters must preclude any conceivable means of working with a smaller number m of parameters ($m < n$) that could specify *all* the group elements $\hat{R}$ of $(\mathcal{G},\star)$.

Let us then denote the elements of the group by $\hat{R}(r_1,r_2,\ldots,r_n)$, or $\hat{R}(\boldsymbol{\rho})$ for short, with $\boldsymbol{\rho} = \{r_1,r_2,\ldots,r_n\}$. It is common to parametrize infinite continuous groups in such a way that the identity element $\hat{E}$ is characterized by vanishing parameters as $\hat{R}(\mathbf{0}) = \hat{R}(0,0,\ldots,0)$.

The group $(\mathcal{G},\star)$ is said to be an *n-parameter* group, or to be a continuous group of *order n*. (Note that although the order of a finite group refers to the number of elements, the order of infinite continuous groups always refers to the number of parameters.) Some books on group theory prefer to speak of the *dimension* of the group, rather than the order.

Notice also that the introduction of continuously varying parameters enables us to speak about the *closeness* or *proximity* between different group elements. A rotation over an angle ω, for example, denoted $\hat{R}(\omega)$, is close to the rotation $\hat{R}(\omega+\epsilon)$ if ϵ is small.

4.1.3 Examples of continuous groups

Because of the central role infinite continuous groups will play in this book, we proceed with a description of four different continuous groups and their properties:

1. The simplest possible example of a continuous one-parameter group is the set of real numbers $\mathbb{R}$ under addition (see §2.4). The elements x of $(\mathbb{R},+)$ are nondenumerable, and the parameter of the group is just the real number x itself, which varies continuously over the interval $]-\infty,+\infty[$.
2. Consider the linear transformation from x to x', representing the rescaling of the real line segment x by a factor a, being a nonzero real constant:

$$x' = ax, \qquad \forall a \in \mathbb{R}_0, \tag{4.1}$$

and let $T(a)$ denote the operator that acts according to the previous transformation. That is,

$$T(a)\,x = x' = ax. \tag{4.2}$$

The set of transformations $\{T(a)\}$ is said to form an infinite continuous group of order 1, with a as the single continuous parameter (Figure 4.2). We can verify this assertion by checking the four group axioms introduced in Chapter 2: closure, associativity, identity, and invertibility. As a result of the closure property of continuous groups, the successive application of two transformations, $T(a_1)$ and $T(a_2)$, should always result in a third transformation $T(a_3)$:

$$T(a_3) = T(a_2)\,T(a_1). \tag{4.3}$$

Note that a_1, a_2, and a_3 do *not* represent three different parameters, but are three possible values of the same parameter a. Now, since

$$\begin{aligned} T(a_2)\,T(a_1)\,x &= T(a_2)\,a_1 x \\ &= a_2 a_1 x, \end{aligned} \tag{4.4}$$

and $T(a_3)\,x = a_3 x$, it follows logically that

$$a_3 = f(a_2,a_1) = a_2 a_1. \tag{4.5}$$

FIGURE 4.2 Having followed the White Rabbit down the rabbit hole, Alice discovers a mysterious bottle, labeled "Drink Me," as well as a piece of cake, saying "Eat Me." The liquid inside the bottle makes her shrink whereas the cake makes her grow. If we represent Alice's length by x, then the amount she drinks corresponds to the operator $T(a)$ with $0 < a \leq 1$, and the amount she eats corresponds to $T(a)$ with $a \geq 1$. Her new length is then given by $x' = T(a)\,x = ax$.

The function $f(a_2, a_1)$ is called a *composition function*. Given any two values a_1 and $a_2 \in \mathbb{R}_0$ characterizing the elements $T(a_1)$ and $T(a_2)$, respectively, the function $f(a_2, a_1)$ tells us how to compute the third value, a_3, which specifies the product element $T(a_3)$ of $T(a_1)$ and $T(a_2)$. Composition functions of infinite continuous groups are thus analogous to the multiplication tables for finite groups (introduced in §2.4) because they both define the structure of the whole group.

The remaining three group postulates can now be examined with the help of $f(a_2, a_1)$. We first note that the composition function in Eq. (4.5) corresponds to the multiplication of two real numbers a_1 and a_2. Since multiplication in $\mathbb{R}_0$ is commutative, the group is clearly Abelian:

$$\begin{gathered} T(a_2)\,T(a_1) = T(a_1)\,T(a_2); \\ a_2\,a_1 = a_1\,a_2. \end{gathered} \tag{4.6}$$

The associative property of multiplication in $\mathbb{R}_0$, in its turn, helps to explain why the associative law holds true:

$$\begin{gathered} T(a_3)\,[T(a_2)\,T(a_1)] = [T(a_3)\,T(a_2)]\,T(a_1); \\ a_3\,[a_2\,a_1] = [a_3\,a_2]\,a_1. \end{gathered} \tag{4.7}$$

Suppose $T(e)$ represents the unit element of the group. It then follows that

$$T(e)\,T(a_1) = T(a_1)\,T(e) = T(a_1). \tag{4.8}$$

Based on the composition function in Eq. (4.5), this translates into $ea_1 = a_1 e = a_1$, which is valid for $e = 1$. It thus follows that $T(1)$ is the identity element. Indeed,

$$T(1)\,x = 1x = x. \tag{4.9}$$

Finally, denoting the inverse element $[T(a_1)]^{-1}$ of $T(a_1)$ by the symbol $T\left(a_1'\right)$, we have

$$\begin{aligned} T(a_1)\,T\left(a_1'\right) = T\left(a_1'\right)T(a_1) = T(1);\\ a_1 a_1' = a_1' a_1 = 1. \end{aligned} \tag{4.10}$$

This holds true for $a_1' = a_1^{-1}$; that is, $[T(a_1)]^{-1} = T\left(a_1^{-1}\right)$. The restriction $a \neq 0$ in Eq. (4.1) follows from the fact that the inverse of 0 (i.e., $1/0$) is undefined. This concludes our proof that the set of transformations $T(a)$ forms an infinite, continuous one-parameter group.

3. Now consider the linear transformation from x to x' corresponding to a rescaling of the real line segment x by a factor a, followed by a translation over b:

$$x' = ax + b, \qquad \forall a \in \mathbb{R}_0 \text{ and } b \in \mathbb{R}. \tag{4.11}$$

Letting $T(a,b)$ denote the operator acting according to Eq. (4.11), we have

$$T(a,b)\,x = x' = ax + b. \tag{4.12}$$

The set of transformations $\{T(a,b)\}$ is said to form a two-parameter group, with a and b as continuous parameters. Because of closure, the product of two transformations $T(a_1,b_1)$ and $T(a_2,b_2)$ always yields a third transformation—say, $T(a_3,b_3)$. By writing

$$\begin{aligned} T(a_2,b_2)\,T(a_1,b_1)\,x &= T(a_2,b_2)\,(a_1 x + b_1)\\ &= a_2\,(a_1 x + b_1) + b_2\\ &= a_2 a_1 x + a_2 b_1 + b_2\\ &= T(a_2 a_1, a_2 b_1 + b_2)\,x, \end{aligned} \tag{4.13}$$

we see that $T(a_3,b_3) = T(a_2 a_1, a_2 b_1 + b_2)$. We thus obtain the following two equalities:

$$a_3 = f_1(a_2,a_1,b_2,b_1) = a_2 a_1; \tag{4.14}$$

$$b_3 = f_2(a_2,a_1,b_2,b_1) = a_2 b_1 + b_2. \tag{4.15}$$

These composition functions should be satisfied at all times to reconcile with the first group postulate. With the help of Eq. (4.14) and Eq. (4.15), the associative law can be shown to hold true. That is,

$$\begin{aligned} T(a_3,b_3)\,[T(a_2,b_2)\,T(a_1,b_1)] = [T(a_3,b_3)\,T(a_2,b_2)]\,T(a_1,b_1);\\ T(a_3,b_3)\,T(a_2 a_1, a_2 b_1 + b_2) = T(a_3 a_2, a_3 b_2 + b_3)\,T(a_1,b_1);\\ T(a_3 a_2 a_1, a_3 a_2 b_1 + a_3 b_2 + b_3) = T(a_3 a_2 a_1, a_3 a_2 b_1 + a_3 b_2 + b_3). \end{aligned} \tag{4.16}$$

Note, however, that this continuous group is not Abelian. After all,

$$\begin{aligned} T(a_2,b_2)\,T(a_1,b_1) \neq T(a_1,b_1)\,T(a_2,b_2);\\ T(a_2 a_1, a_2 b_1 + b_2) \neq T(a_1 a_2, a_1 b_2 + b_1). \end{aligned} \tag{4.17}$$

The identity element clearly corresponds to $T(1,0)$ since

$$T(1,0)\,x = 1x + 0 = x. \tag{4.18}$$

Finally, denoting the inverse of $T(a_1,b_1)$ by the symbol $T\left(a_1',b_1'\right)$, we have

$$T(a_1,b_1)\,T\left(a_1',b_1'\right) = T\left(a_1',b_1'\right)T(a_1,b_1) = T(1,0). \tag{4.19}$$

This means that

$$a_1 a_1' = a_1' a_1 = 1; \quad a_1 b_1' + b_1 = a_1' b_1 + b_1' = 0, \tag{4.20}$$

from which one obtains $a_1' = 1/a_1$ and $b_1' = -b_1/a_1$, thus identifying the inverse element as $\hat{T}\left(1/a_1, -b_1/a_1\right)$. The existence of an inverse element explains once again the requirement that $a \neq 0$.

4. The rotational symmetry group of the circle is continuous as well. In this case, the arbitrary rotation angle ω represents the parameter that varies continuously over the finite interval $[0,\ 2\pi]$, and that can be used to specify the group elements, as expressed by the notation $\hat{R}(\omega)$ in the previous chapter. Because the group elements depend on the rotation angle only, the axial rotation group is a one-parameter group.

   The composition of two rotations through two arbitrary angles ω_1 and ω_2 corresponds to a single rotation over the angle $\omega_1 + \omega_2$:

$$\hat{R}(\omega_2)\,\hat{R}(\omega_1) = \hat{R}\big(f(\omega_2,\omega_1)\big) = \hat{R}(\omega_1+\omega_2) = \hat{R}(\omega_3). \tag{4.21}$$

   The composition function that describes the product rule is thus given by

$$f(\omega_2,\omega_1) = \omega_1 + \omega_2 = \omega_3. \tag{4.22}$$

   This function can be used subsequently to prove the remaining group axioms and the Abelian character of the axial rotation group. The axial rotation group is isomorphic to the matrix group of special orthogonal 2 × 2 matrices, denoted SO(2), as illustrated in the previous chapter.

4.1.4 The composition functions

Generalizing from the previous examples, the closure property of infinite continuous groups is seen to give rise to a set of n continuous *composition functions* $f_i(\boldsymbol{\sigma};\boldsymbol{\rho})$ $(i = 1,\ldots,n)$ from which the parameters $\boldsymbol{\tau} = \{t_1,t_2,\ldots,t_n\}$ of any particular product element $\hat{T} = \hat{S}\hat{R}$ can be calculated, given the parameters $\boldsymbol{\rho} = \{r_1,r_2,\ldots,r_n\}$ and $\boldsymbol{\sigma} = \{s_1,s_2,\ldots,s_n\}$ of the factor elements $\hat{R}$ and $\hat{S}$, respectively:

$$t_i = f_i(s_1,s_2,\ldots,s_n;r_1,r_2,\ldots,r_n) = f_i(\boldsymbol{\sigma};\boldsymbol{\rho}), \qquad i = 1,\ldots,n. \tag{4.23}$$

The composition functions $f_i(\boldsymbol{\sigma};\boldsymbol{\rho})$ define the structure of the whole group in much the same way as the multiplication table does for finite groups.

To qualify as a group, the remaining three group postulates must be satisfied. These group axioms impose a number of additional constraints on the composition functions:

$$\begin{aligned}
&\textit{Associativity:} && f_i\big(\boldsymbol{\tau};\boldsymbol{f}(\boldsymbol{\sigma};\boldsymbol{\rho})\big) = f_i\big(\boldsymbol{f}(\boldsymbol{\tau};\boldsymbol{\sigma});\boldsymbol{\rho}\big);\\
&\textit{Identity:} && f_i(\boldsymbol{\varepsilon};\boldsymbol{\rho}) = f_i(\boldsymbol{\rho};\boldsymbol{\varepsilon}) = r_i;\\
&\textit{Invertibility:} && f_i\left(\boldsymbol{\rho}^{-1};\boldsymbol{\rho}\right) = f_i\left(\boldsymbol{\rho};\boldsymbol{\rho}^{-1}\right) = e_i,
\end{aligned} \tag{4.24}$$

where $\boldsymbol{\varepsilon} = \{e_1, e_2, \ldots, e_n\}$ are the parameters of the identity element $\hat{E}$ and $\boldsymbol{\rho}^{-1} = \left\{r_1^{-1}, r_2^{-1}, \ldots, r_n^{-1}\right\}$ are the parameters of the inverse element $\hat{R}^{-1}$ of $\hat{R}$.

4.2 LIE GROUPS

With the introduction of the composition functions $f_i(\boldsymbol{\sigma};\boldsymbol{\rho})$ of an infinite continuous group in §4.1.4, we are now in a position to provide a definition of *Lie groups* in §4.2.1. We will then study the *parameter space* of a Lie group in §4.2.2, as well as the properties of *connectedness* and *compactness* in §4.2.3, before we move on to the central concept of Lie group theory—the *infinitesimal generator*—in §4.3.

4.2.1 Definition

To call an infinite continuous group a Lie group, there is one more requirement of the composition functions $f_i(\boldsymbol{\sigma};\boldsymbol{\rho})$ that should be satisfied: the parameter $\boldsymbol{\tau}$ of the product element $\hat{T} = \hat{S}\hat{R}$ should be an *analytic function* of the parameters $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ of the factor elements $\hat{R}$ and $\hat{S}$.

Definition 4.1 (Lie group): Let $(\mathcal{G}, \star)$ be an infinite continuous group of order n, with a product rule that is characterized by n composition functions $f_i(\boldsymbol{\sigma};\boldsymbol{\rho})$ $(i = 1, \ldots, n)$. If the composition functions are *analytic functions*, then the corresponding infinite continuous group is called an n-parameter *Lie group*. ■

With this simple qualification the whole arsenal of mathematical analysis at once enters the picture. Recall that a function $f(x)$ is said to be *analytic* when it can be written as a power series:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots . \tag{4.25}$$

As a consequence, analytic functions are *smooth* (*infinitely differentiable*) functions,[7] which can be written as a *Taylor series expansion* around any point in their domain (see Appendix C).

Notice that the composition functions introduced in the previous examples (Eqs. (4.5), (4.14), (4.15), and (4.22)) are analytic in this sense. Hence, all the infinite continuous groups in §4.1.3 are also Lie groups.

Lie groups represent the best-developed theory of continuous symmetry of mathematical objects and structures. This makes them indispensable tools for many parts of contemporary mathematics, as well as for modern theoretical physics and chemistry.

---

[7] A function f is said to be of class C^k (where k is a nonnegative integer) when its derivatives $f', f'', \ldots, f^k$ exist and are all continuous. The class C^0 thus consists of all continuous functions. Functions whose first derivative is continuous as well belong to the class C^1, and so forth. A function f is said to be of class C^∞ if it has derivatives of all orders that are all continuous. The function f is then said to be *smooth* or *infinitely differentiable*. Moreover, a smooth function f is said to be of class C^ω, or to be *analytic*, if it can be written as a *Taylor series expansion* around any point in its domain (see also Appendix C).

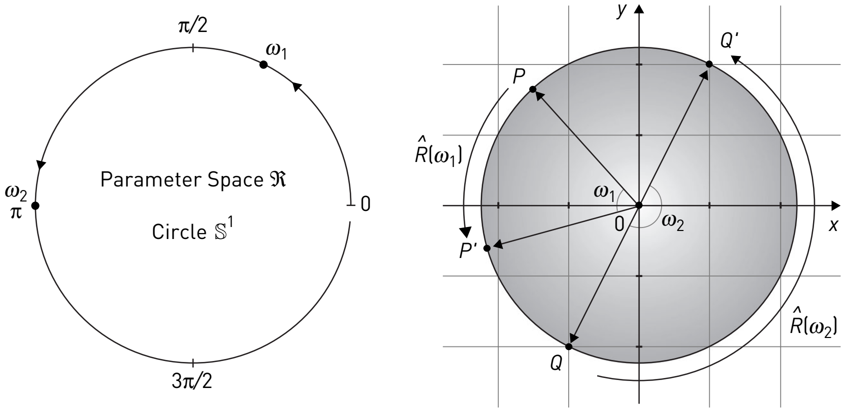


FIGURE 4.3 The axial rotation group is a one-parameter group, whose only parameter, the rotation angle ω, varies over the closed and bounded interval $[0, 2\pi]$. As a result, the parameter space $\Re$ corresponds to the one-dimensional real line segment, where ω varies from 0 to 2π. What is more, because a rotation over 2π corresponds to the null rotation, the end points 0 and 2π should be brought into correspondence with one another, which results in the formation of a unit circle $\mathbb{S}^1$. To each point ω in the one-dimensional parameter space $\mathbb{S}^1$ then corresponds a physical rotation in two-dimensional space over an angle ω, as specified by the rotation operator $\hat{R}(\omega)$. This is illustrated in the figure for two different angles: ω_1 and ω_2.

4.2.2 Parameter space

The parameters r_i $(i = 1, \ldots, n)$ that label the elements of an infinite continuous Lie group $(\mathcal{G}, \star)$ vary in an n-dimensional topological space, which is called the *parameter space* $\Re_n$ of $(\mathcal{G}, \star)$ (also known variously as the *group space* or *group manifold*). Each point in this parameter space represents a particular set of n parameter values r_i; the parameters r_i are said to form the components of an n-dimensional vector $\boldsymbol{\rho} = (r_1, r_2, \ldots, r_n)$.

For each vector $\boldsymbol{\rho}$ in $\Re_n$, there is a corresponding element $\hat{R}(\boldsymbol{\rho})$ in the group $(\mathcal{G}, \star)$. The identity element $\hat{R}(\mathbf{0}) = \hat{R}(0, 0, \ldots, 0)$, for example, is mapped onto the origin of the parameter space $\Re_n$. It would be tempting to claim that this mapping is one to one—that is, that the parameter space $\Re_n$ is isomorphic to the group $(\mathcal{G}, \star)$. This, however, is usually not the case.

Consider the SO(2) group as a case in point. Each rotation operator $\hat{R}(\omega)$ is labeled by one parameter, the rotation angle ω. We might thus be led to believe that the parameter space is the one-dimensional real line, where ω varies from $-\infty$ to $+\infty$. However, because each rotation over an angle ω outside the interval $[0, 2\pi]$ is brought back to a rotation over an angle $\omega \in [0, 2\pi]$ via the equation

$$\hat{R}(\omega) = \hat{R}(\omega + 2\pi n), \qquad \forall n \in \mathbb{Z}, \tag{4.26}$$

the parameter space actually corresponds to the one-dimensional space defined by all points on the unit circle $\mathbb{S}^1 \subset \mathbb{R}^2$.[8] To each point ω in the parameter space $\mathbb{S}^1$ corresponds a rotation $\hat{R}(\omega)$ over an angle ω (Figure 4.3). The circle defines the *topology* of the parameter space (see also Figure 4.4).

---

[8] This notation is explained in Definition 9.6 in Chapter 9.

FIGURE 4.4 "'What size do you want to be?' asked the Caterpillar. 'Well, I should like to be a little larger, sir, if you wouldn't mind,' said Alice. The Caterpillar took the hookah out of its mouth and remarked: 'One side [of the mushroom] will make you grow taller, and the other side will make you grow shorter.' " "Alice remained looking thoughtfully at the mushroom for a minute, trying to make out which were the two sides of it; and as it was perfectly round, she found this a very difficult question. However, at last she stretched her arms round it as far as they could go, and broke off a bit of the edge with each hand." The mushroom represents the parameter space $\Re$, with parameter ω that varies continuously as Alice goes round the mushroom, and determines to what extent she grows or shrinks. Nibbling a bit from one piece, has the physical effect of making her shrink, whereas eating a little from the other piece makes her grow.

Finally, we note that two elements $\hat{R}(\boldsymbol{\rho})$ and $\hat{R}(\boldsymbol{\sigma})$ are said to be "close" to one another if the distance $\left[\sum_{i=1}^{n}(r_i - s_i)^2\right]^{1/2}$ between the corresponding points $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ is small.

4.2.3 Connectedness and compactness

Let $(\mathcal{G},\star)$ represent an infinite continuous group of order n, with corresponding parameter space $\Re_n$, and consider any two group elements $\hat{R}(\boldsymbol{\rho})$ and $\hat{R}(\boldsymbol{\sigma})$ with corresponding points $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ in $\Re_n$. If $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ can be "connected" with one another via one or more paths that lie entirely within the parameter space, then $\Re_n$ is said to be *continuously connected*.

Each point on this path in parameter space then corresponds to a group element, the entire set of which forms another path that connects the two group elements $\hat{R}(\boldsymbol{\rho})$ and $\hat{R}(\boldsymbol{\sigma})$. An infinite continuous group $(\mathcal{G},\star)$ is thus said to be *connected* when any two arbitrary elements from the group can be connected with one another via a continuous variation of the group parameters (i.e., when its parameter space is connected).

The SO(2) group is a connected group in this sense because any rotation $\hat{R}(\omega_1)$ can be turned into any other rotation $\hat{R}(\omega_2)$ via a continuous variation of the rotation angle from ω_1 to ω_2. The O(2) group of rotations and reflections, on the other hand, forms a *disconnected* group. To see this, let us consider the identity transformation $\hat{E}=\hat{R}(0)$ or, in matrix form, $\mathbb{R}(0)=\mathbb{I}$. Its determinant equals

$$\det\mathbb{R}(0)=\det\mathbb{I}=1. \tag{4.27}$$

Every operator $\hat{R}(\omega)$ that evolves continuously from $\hat{E}$ must therefore satisfy the condition

$$\det\mathbb{R}(\omega)=1, \qquad \forall\omega\in[0,2\pi]. \tag{4.28}$$

That is, every continuous path starting from $\hat{E}$ connects operators with unit determinant, $\det=+1$. These are *proper* rotations. All *improper* rotations (or reflections), on the other hand, with $\det=-1$, are disconnected from the identity operator.

Topologically, the parameter space of the orthogonal group O(2) thus consists of two disconnected circles. One circle corresponds to the parameter space of the SO(2) subgroup and contains all two-dimensional rotations, including the identity transformation. The other circle contains all rotation reflections, consisting of all rotations followed by the inversion operator.

To conclude, an infinite continuous group $(\mathcal{G},\star)$ of order n is said to be *compact* when its corresponding parameter space $\mathfrak{R}_n$ is compact in the sense of being a closed and bounded space. This means the domain of variation for the different parameters should be finite. Both the O(2) and SO(2) groups are compact because the rotation angle varies over the finite interval $[0,2\pi]$. In the third part of this book, we will encounter noncompact groups, such as the SO(2, 1) and SO(4, 2) groups, where some of the parameters vary between $-\infty$ and $+\infty$.

4.3 THE INFINITESIMAL GENERATOR

So far, the identity transformation $\hat{E}$ of the axial rotation group has been parametrized as $\hat{R}(0)$, with $\omega=0$. Infinitesimally small values of the rotation angle, denoted $\delta\omega$, are therefore associated with elements that are infinitesimally close to the identity element, and they correspond to infinitesimally small rotations, denoted $\hat{R}(\delta\omega)$.

Now, because every finite rotation $\hat{R}(\omega)$ is *continuously connected* to the identity operator $\hat{E}$, one can produce any desired finite rotation $\hat{R}(\omega)$ by the repeated application (or integration) of the infinitesimal rotation $\hat{R}(\delta\omega)$:

$$\hat{R}(0)\rightarrow\hat{R}(\delta\omega)\rightarrow\hat{R}(\delta\omega+\delta\omega)\rightarrow\ldots\rightarrow\hat{R}(\omega), \tag{4.29}$$

where use was made of the fact that $\hat{R}(\delta\omega)\,\hat{R}(\delta\omega) = \hat{R}(\delta\omega+\delta\omega)$.[9] Thus, it proves especially interesting to study the nature of the group elements *close to the identity*.

This study will lead us to the concept of the *infinitesimal generator*, which was first introduced by Sophus Lie. In the following two sections, §4.3.1 and §4.3.2, the generator of the SO(2) group will be derived in both matrix and operator form. The great advantage behind the introduction of this concept is that we no longer have to consider the group as a whole, with its *infinite* number of elements; we can limit our attention to a study of the *finite* set of infinitesimal transformations around the identity—the underlying reason being that a perfectly symmetric object, such as a circle, is uniform everywhere, which makes it sufficient to probe its local curvature at only one point to grasp the structure of the whole.

4.3.1 Matrix form of the SO(2) generator

Because the principle of expanding a function in Taylor and MacLaurin series is used extensively in the following paragraphs, refer to Appendix C for a brief review of this powerful mathematical tool. Now, to derive the *infinitesimal generator* of the continuous Lie group SO(2), we start by expanding the matrix $\mathbb{R}(\omega)$ in a Taylor series around the identity ($\omega = 0$). This Maclaurin series can be written with the help of Eq. (C.16):

$$\mathbb{R}(\omega) = \mathbb{R}(0) + \left.\frac{d\mathbb{R}(\omega)}{d\omega}\right|_{\omega=0} \omega + \frac{1}{2!}\left.\frac{d^2\mathbb{R}(\omega)}{d\omega^2}\right|_{\omega=0} \omega^2 + \dots . \tag{4.30}$$

The expansion coefficients in this series can be determined directly from the matrix form of $\mathbb{R}(\omega)$:

$$\mathbb{R}(\omega) = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}, \tag{4.31}$$

by differentiating each matrix element an appropriate number of times. However, because there are an infinite number of coefficients to calculate, it proves more interesting to substitute this brute-force method by the following, more elegant, method.

We start by determining the first coefficient of Eq. (4.30) by taking the first derivative of $\mathbb{R}(\omega)$ at $\omega = 0$:

$$\begin{aligned} \left.\frac{d\mathbb{R}(\omega)}{d\omega}\right|_{\omega=0} &= \left.\begin{bmatrix} d\cos\omega/d\omega & -d\sin\omega/d\omega \\ d\sin\omega/d\omega & d\cos\omega/d\omega \end{bmatrix}\right|_{\omega=0} \\ &= \left.\begin{bmatrix} -\sin\omega & -\cos\omega \\ \cos\omega & -\sin\omega \end{bmatrix}\right|_{\omega=0} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \equiv \mathbb{X}. \end{aligned} \tag{4.32}$$

To determine the higher order derivatives of $\mathbb{R}(\omega)$, we continue by noting that each rotation $\mathbb{R}(\omega)$ can be written as $\mathbb{R}(\omega_1+\omega_2)$, where we will later set $\omega_1 = 0$ and

---

[9] It bears repeating, once again, that not every group element needs to be accessible in this sense from the repeated application of the elements close to the identity. In the O(2) group, for instance, the elements with determinant -1 cannot be reached from the identity with determinant $+1$. See §4.2.3 for more information.

$\omega_2 = \omega$. For the moment, we can use Eq. (4.21) to express the derivative of $\mathbb{R}(\omega_1 + \omega_2)$ with respect to ω_1 as

$$\frac{d\mathbb{R}(\omega_1+\omega_2)}{d\omega_1} = \frac{d\mathbb{R}(\omega_1)}{d\omega_1}\mathbb{R}(\omega_2). \tag{4.33}$$

With the help of the chain rule, the left-hand side of this equation can be rewritten as

$$\frac{d\mathbb{R}(\omega_1+\omega_2)}{d\omega_1} = \frac{d\mathbb{R}(\omega_1+\omega_2)}{d(\omega_1+\omega_2)}\frac{d(\omega_1+\omega_2)}{d\omega_1}. \tag{4.34}$$

Setting $\omega_1 = 0$,

$$\left[\frac{d\mathbb{R}(\omega_1+\omega_2)}{d(\omega_1+\omega_2)}\frac{d(\omega_1+\omega_2)}{d\omega_1}\right]\Bigg|_{\omega_1=0} = \frac{d\mathbb{R}(\omega_2)}{d\omega_2}. \tag{4.35}$$

This result can be plugged into Eq. (4.33) to yield

$$\frac{d\mathbb{R}(\omega)}{d\omega} = \frac{d\mathbb{R}(\omega_1)}{d\omega_1}\bigg|_{\omega_1=0}\mathbb{R}(\omega) = \mathbb{X}\mathbb{R}(\omega), \tag{4.36}$$

where we have used the result from Eq. (4.32) and where ω_2 has been replaced by ω. The higher order derivatives of $\mathbb{R}(\omega)$ can then be written as

$$\frac{d^n\mathbb{R}(\omega)}{d\omega^n}\bigg|_{\omega=0} = \mathbb{X}\,\frac{d^{n-1}\mathbb{R}(\omega)}{d\omega^{n-1}}\bigg|_{\omega=0}. \tag{4.37}$$

And, with the help of Eq. (4.32),

$$\frac{d^n\mathbb{R}(\omega)}{d\omega^n}\bigg|_{\omega=0} = \mathbb{X}^n. \tag{4.38}$$

Substituting this result into the Maclaurin series in Eq. (4.30) finally yields

$$\mathbb{R}(\omega) = \mathbb{I} + \mathbb{X}\omega + \frac{1}{2!}\mathbb{X}^2\omega^2 + \frac{1}{3!}\mathbb{X}^3\omega^3 + \dots, \tag{4.39}$$

which can be written in the more compact sigma notation

$$\mathbb{R}(\omega) = \sum_{n=0}^{\infty}\frac{1}{n!}(\mathbb{X}\omega)^n, \tag{4.40}$$

where $\mathbb{X}^0 = \mathbb{I}$. Noting the similarity of this equation to Eq. (C.15), it follows that

$$\mathbb{R}(\omega) = e^{\omega\mathbb{X}}. \tag{4.41}$$

That is to say, every proper two-dimensional rotation $\mathbb{R}(\omega)$ by a finite angle ω can be obtained from the exponentiation of the matrix $\mathbb{X}$, which is therefore called the *generator* of the axial rotation group SO(2). Because $\mathbb{X}$ is a matrix, this is called the *matrix form* of the generator.

The expression in Eq. (4.41) is commonly referred to as the *exponential map* because it forms the bridge between the generator $\mathbb{X}$ and the group elements $\mathbb{R}(\omega)$. We will come back to this important notion in the following chapters.

Notice also that the Taylor expansion in Eq. (4.39) can be *approximated* by ignoring all the terms of order ω^2 or higher:

$$\begin{aligned} \mathbb{R}(\omega) &= \mathbb{I} + \mathbb{X}\omega + \mathcal{O}\left(\omega^2\right); \\ \mathbb{R}(\omega) &\approx \mathbb{I} + \mathbb{X}\omega. \end{aligned} \tag{4.42}$$

This expression becomes *exact* when infinitesimally small rotations are considered. We can thus write

$$\mathbb{R}(\delta\omega) = \mathbb{I} + \mathbb{X}\delta\omega. \tag{4.43}$$

This clearly shows that the infinitesimal generator $\mathbb{X}$ is generating infinitesimally small rotations $\hat{R}(\delta\omega)$.

In summary, we saw that most of the SO(2) group structure is determined by the generator $\mathbb{X}$, which in turn is determined by the *local behavior* of the SO(2) group around the identity. The *global behavior* of the SO(2) group cannot be captured by the generator concept, however. These are mostly properties of a topological nature, such as the relation in Eq. (4.26).

4.3.2 Operator form of the SO(2) generator

An alternative way of representing the infinitesimal operator $\mathbb{X}$ of the Lie group SO(2) is in terms of a *differential operator*. This so-called *operator form* of the generator will be most helpful in the next section, when a connection will be made with quantum mechanics. To derive the operator associated with an infinitesimal rotation, we start by considering a rotation over a small angle $\Delta\omega$, denoted $\hat{R}(\Delta\omega)$. In line with the defining relationship in Chapter 3, Eq. (3.40), this rotates the angular coordinate of all points backward over an angle $\Delta\omega$:

$$\begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{R}(\Delta\omega) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \cos\Delta\omega & -\sin\Delta\omega \\ \sin\Delta\omega & \cos\Delta\omega \end{bmatrix}. \tag{4.44}$$

In the limit where the angle becomes infinitesimally small ($\Delta\omega \to 0$), this equation becomes

$$\begin{aligned} \begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{R}(\delta\omega) &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & -\delta\omega \\ \delta\omega & 1 \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} (\mathbb{I} + \mathbb{X}\delta\omega). \end{aligned} \tag{4.45}$$

Thus, $x' = x + y\delta\omega$ and $y' = y - x\delta\omega$. Based on this fact, we can reproduce these results by a differential operator as follows:

$$\hat{R}(\delta\omega) \begin{bmatrix} x & y \end{bmatrix} = \left(\hat{R}(0) + \hat{X}\delta\omega \right) \begin{bmatrix} x & y \end{bmatrix}. \tag{4.46}$$

The operator $\hat{X}$ is defined in terms of the rotation operators as

$$\hat{X} = \lim_{\delta\omega \to 0} \frac{\hat{R}(\delta\omega) - \hat{R}(0)}{\delta\omega}. \tag{4.47}$$

This operator is associated with the matrix transformation $\mathbb{X}$ and can thus be written as

$$\hat{X} \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{X}. \tag{4.48}$$

To obtain the algebraic form of the operator, we first note that the product of the partial derivatives arranged in a column with the row matrix of the coordinative functions is the unit matrix

$$\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix} = \mathbb{I}. \tag{4.49}$$

Introducing this product to the right of Eq. (4.48) yields

$$\hat{X} \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{X} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix}. \tag{4.50}$$

The operator itself can thus be written as

$$\hat{X} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{X} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix}, \tag{4.51}$$

or in algebraic form as

$$\hat{X} = \left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \right). \tag{4.52}$$

As outlined in the following section, this expression for the generator of the SO(2) rotation group is closely related to the z-component of the *angular momentum operator*. Because the SO(2) group is the simplest of all Lie groups, the full benefits of the infinitesimal generator are not yet readily apparent. However, we will see in our examination of the SO(3) group in Chapter 5 that infinitesimal generators embody much of the structure of the full group. In fact, most solutions to the quantum mechanical problems in this book are found by using the infinitesimal operators *only*, thereby disregarding all other operators completely.

4.4 ANGULAR MOMENTUM

4.4.1 Classical mechanical picture

There is no better way to illustrate the importance of the generator we just derived than to switch to the physical treatment of a rotational motion around a fixed point. Classical physics first. A planet orbiting around the sun unknowingly conserves two important physical quantities: its *energy* and its *angular momentum*. The angular momentum vector $\mathbf{L}$ of a moving particle of mass m in three-dimensional space is defined as follows. Let $\mathbf{r}$ be the vector from the origin to the instantaneous position of the particle. We have

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z, \tag{4.53}$$

where x, y, and z are the particle's coordinates at a given instant. These coordinates are functions of time. Defining the velocity vector $\mathbf{v}$ as the time derivative of the position vector, we have

$$\begin{aligned} \mathbf{v} \equiv \frac{d\mathbf{r}}{dt} &= \frac{dx}{dt}\mathbf{e}_x + \frac{dy}{dt}\mathbf{e}_y + \frac{dz}{dt}\mathbf{e}_z; \\ v_x &= \frac{dx}{dt},\ v_y = \frac{dy}{dt},\ v_z = \frac{dz}{dt}. \end{aligned} \tag{4.54}$$

We define the particle's *linear momentum vector* $\mathbf{p}$ by

$$\begin{aligned} \mathbf{p} &\equiv m\mathbf{v}; \\ p_x = mv_x,\ p_y &= mv_y,\ p_z = mv_z. \end{aligned} \tag{4.55}$$

The time derivative of the velocity is the acceleration $\mathbf{a}$. By Newton's law, the product of mass and acceleration corresponds to the force (i.e., $\mathbf{F} = m\mathbf{a}$). For a particle moving in a central field, the force is directed along the radius; hence,

$$\mathbf{F} = F\mathbf{e}_r = F\frac{\mathbf{r}}{r}, \tag{4.56}$$

with the direction given by the unit vector $\mathbf{e}_r = \mathbf{r}/r$. The particle's *angular momentum* $\mathbf{L}$ is defined in classical mechanics as the *outer product* (see Appendix B) of the particle's position and momentum:

$$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p};$$

$$\mathbf{L} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} = \begin{vmatrix} y & z \\ p_y & p_z \end{vmatrix} \mathbf{e}_x - \begin{vmatrix} x & z \\ p_x & p_z \end{vmatrix} \mathbf{e}_y + \begin{vmatrix} x & y \\ p_x & p_y \end{vmatrix} \mathbf{e}_z; \tag{4.57}$$

$$L_x = yp_z - zp_y,\ L_y = zp_x - xp_z,\ L_z = xp_y - yp_x.$$

L_x, L_y, and L_z are the components of **L** along the x-, y-, and z-axes. The angular momentum vector **L** is perpendicular to the plane defined by the particle's position vector **r** and its velocity **v**. The time evolution of the angular momentum is given by

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \frac{d\mathbf{r}}{dt} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt} \\ &= \mathbf{v} \times \mathbf{p} + \mathbf{r} \times m\mathbf{a} \\ &= \frac{1}{m}\mathbf{p} \times \mathbf{p} + \frac{F}{r}\mathbf{r} \times \mathbf{r} = \mathbf{0}. \end{aligned} \tag{4.58}$$

A particle orbiting around a fixed point is attracted by a central force (i.e., a force directed along the radius from the origin). As a result, both vector products in the equation vanish and the angular momentum is conserved. This means the angular momentum is a constant that does not change with time. When the orbital plane coincides with the xy-plane, the angular momentum vector is directed along the z-axis. The only nonzero angular momentum is, then, the L_z component.

4.4.2 Quantum mechanical picture

Now for quantum mechanics.[10] In quantum mechanics, the analogue of the classical quantity **L** is the so-called *orbital angular momentum*, which results from the motion of a particle through space. In addition, there is a *spin angular momentum* that is an intrinsic property of elementary particles and has no classical mechanical analogue (see Chapter 8). For the moment, let us restrict our exploration to a quantum mechanical consideration of the orbital angular momentum.

In quantum mechanics, the physical quantities become operators. We get the proper form of the angular momentum operators by replacing the coordinates and momenta in the classical Eq. (4.57) with their corresponding operators according to Eqs. (D.19) and (D.20): $\hat{x} = x$ and $\hat{p}_x = \hbar/i\ \partial/\partial x = -i\hbar\partial/\partial x$. Here $\hbar$ denotes $h/2\pi$, with h representing Planck's constant. We find

$$\begin{aligned} \hat{L}_x &= -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right); \\ \hat{L}_y &= -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right); \\ \hat{L}_z &= -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \end{aligned} \tag{4.59}$$

Using Eq. (4.52), the quantum mechanical operator for the z-component of the orbital angular momentum $\hat{L}_z$ becomes

$$\hat{L}_z = i\hbar\hat{X}. \tag{4.60}$$

---

[10] For a brief review of quantum mechanics, see Appendix D.

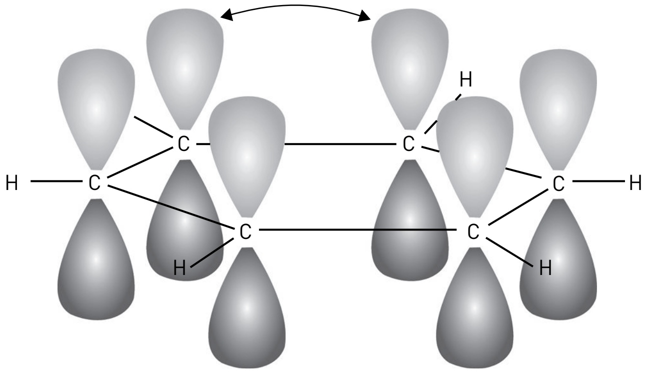


FIGURE 4.5 Benzene structure. Electrons can rotate around by hopping from one p_z-orbital to another on the neighbouring carbon atoms.

It follows that the infinitesimal generator of the group of rotations in the *xy*-plane, multiplied by $i\hbar$, equals the *z*-component of the angular momentum operator $\hat{L}_z$. Let us demonstrate the power of this Lie algebraic concept by solving the wave equation for a particle on a ring.

4.5 SO(2) SYMMETRY AND AROMATIC MOLECULES

Angular momentum, and thus SO(2) symmetry, is a key feature in the electronic structure of circular molecules. This is best illustrated for the particular family of the *annulenes*, consisting of a necklace of methyn (CH) beads, each of which carries an electron in a $2p_z$-orbital, perpendicular to the plane of the ring. The coupling or "conjugation" of these electrons gives rise to a delocalized electronic system, commonly denoted as an $n\pi$ system, where n is the number of valence electrons in the $2p_z$-orbitals. The prototype of annulenes is the 6π-molecule benzene, C_6H_6, which has the shape of a regular hexagon (Figure 4.5). To highlight the cyclic nature of these electronic structures, we now look at the quantum mechanical model of a particle, *in casu* an electron, on a ring. Despite the fact that this is an extreme simplification of the actual molecular systems, this model pinpoints the essence of cyclic electronic structures.

4.5.1 The particle on a ring model

The model consists of a ring surrounded by extremely high-potential walls, so that a particle on the ring can move along the circumference only. Its position in cylindrical coordinates is thus given by a fixed radius r, and a free angular coordinate ϕ, which ranges from 0 to 2π.

Angular momentum

For a ring in the *xy*-plane, the *z*-component of the angular momentum is conserved, as we proved in the previous section. In quantum mechanics, this means the wave function $|\psi\rangle$, which describes the physical state of the moving particle, is an eigenfunction of $\hat{L}_z$:

$$\hat{L}_z|\psi\rangle = \hbar k|\psi\rangle. \tag{4.61}$$

The eigenvalue on the right-hand side of this equation is denoted as $\hbar k$, where $\hbar$ is the unit of angular momentum and k is a dimensionless quantity. The expression for the z-component of the angular momentum operator, given in Eq. (4.59), can be used to rewrite the previous equation:

$$-i\hbar\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)|\psi\rangle=\hbar k|\psi\rangle. \tag{4.62}$$

We now replace the x- and y-coordinates with the angular coordinate ϕ. In the xy-plane the coordinate transformation is as follows:

$$\begin{aligned} x &= r\cos\phi; \\ y &= r\sin\phi. \end{aligned} \tag{4.63}$$

The *multivariable chain rule* allows us to express the derivative with respect to ϕ as

$$\frac{\partial}{\partial\phi}=\frac{\partial x}{\partial\phi}\frac{\partial}{\partial x}+\frac{\partial y}{\partial\phi}\frac{\partial}{\partial y}. \tag{4.64}$$

From Eq. (4.63), we obtain the required derivatives as

$$\begin{aligned} \frac{\partial x}{\partial\phi} &= -r\sin\phi=-y; \\ \frac{\partial y}{\partial\phi} &= r\cos\phi=x. \end{aligned} \tag{4.65}$$

Hence, we find

$$\frac{\partial}{\partial\phi}=-y\frac{\partial}{\partial x}+x\frac{\partial}{\partial y}. \tag{4.66}$$

Substituting this expression in Eq. (4.62) gives

$$-i\hbar\frac{\partial}{\partial\phi}|\psi(\phi)\rangle=\hbar k|\psi(\phi)\rangle. \tag{4.67}$$

Dividing by $\hbar$ and multiplying both sides by i gives

$$\frac{\partial}{\partial\phi}|\psi(\phi)\rangle=ik|\psi(\phi)\rangle. \tag{4.68}$$

This equation is a first-order differential equation, which can be solved easily:

$$|\psi(\phi)\rangle=\frac{1}{N}e^{ik\phi}. \tag{4.69}$$

The value of k is quantized as a result of the cyclic nature of the angular coordinate, which requires that the wave function be continuous and single-valued in the periodic interval. This implies that the value of the wave function must be the same at ϕ and $\phi+2\pi$:

$$\begin{aligned} |\psi(\phi)\rangle &= |\psi(\phi+2\pi)\rangle \\ e^{ik\phi} &= e^{ik\phi}e^{ik2\pi} \\ 1 &= e^{ik2\pi}. \end{aligned} \tag{4.70}$$

The equality in Eq. (4.70) is satisfied only if k is an integer. Integer values of k are traditionally denoted as m_l and are known as *magnetic quantum numbers* when referring to electronic states. The possible values for m_l are as follows:

$$m_l=0,\ \pm1,\ \pm2,\ \pm3,\ \ldots. \tag{4.71}$$

The solution of the differential equation also yields an undetermined coefficient $1/N$, which is called the *normalization constant* of the wave function.[11] This normalization constant is usually determined by requiring that the probability density, which corresponds to the product of the wave function with its complex conjugate, normalizes to 1 when integrated over all space. In our case, the limits of integration for normalization are from 0 to 2π, because this covers the entire circular path.

$$\langle \psi \mid \psi \rangle = \frac{1}{N^2} \int_0^{2\pi} \left(e^{-im_l\phi} \right) \left(e^{im_l\phi} \right) d\phi = \frac{1}{N^2} \int_0^{2\pi} d\phi = \frac{1}{N^2} 2\pi = 1; \qquad (4.72)$$

$$N = \sqrt{2\pi}.$$

The normalized wave function for the particle on a ring becomes

$$|\psi(\phi)\rangle = \sqrt{\frac{1}{2\pi}} e^{im_l\phi}; \quad m_l = 0, \pm 1, \pm 2, \pm 3, \dots . \qquad (4.73)$$

A physical connection can be made with the sign of the m_l quantum number. Positive and negative signs of m_l indicate the direction in which the wave is propagating. Based on the right-hand rule, a positive value for the angular momentum indicates a counterclockwise rotation whereas negative values indicate a clockwise rotation. Because the value for angular momentum is given by $\hbar k = \hbar m_l$ according to Eq. (4.61), counterclockwise rotations correspond to positive values of m_l, and clockwise rotations indicate negative values of m_l.

Energy

As in classical physics, energy is also conserved. The corresponding operator is the Hamiltonian $\hat{\mathscr{H}}$ and the eigenvalue equation for energy is the celebrated Schrödinger equation (see Appendix D):

$$\hat{\mathscr{H}} |\psi\rangle = E|\psi\rangle, \qquad (4.74)$$

where E is the eigenvalue of $\hat{\mathscr{H}}$. The time-independent Schrödinger equation for a single particle of mass m moving along one dimension reads as

$$-\frac{\hbar^2}{2m} \frac{d^2 |\psi(x)\rangle}{dx^2} + V(x) |\psi(x)\rangle = E |\psi(x)\rangle . \qquad (4.75)$$

In the case of a particle on a ring, the coordinate x corresponds to the progression along the perimeter of the ring and is denoted as u. The potential energy anywhere on the circle can be taken as the zero point of energy: $V(u) = 0$. The Hamiltonian, which is the sum of the kinetic and potential energies, then reduces to the kinetic term. The Schrödinger equation thus becomes

$$\hat{\mathscr{H}} |\psi(u)\rangle = -\frac{\hbar^2}{2m} \frac{d^2 |\psi(u)\rangle}{du^2} = E|\psi(u)\rangle. \qquad (4.76)$$

Because this single particle is confined to a circle with constant radius r, the u-coordinate has, in fact, been bent upon itself and runs from 0 to $2\pi r$. We are working, in other words, within a *one-dimensional periodic space.* All u-positions can therefore be rewritten in function of r and ϕ. We find that

$$u(r, \phi) = \phi r, \qquad (4.77)$$

---

[11] Normalization leaves the phase of N undetermined. For simplicity, we take N to be a positive real number.

with ϕ running from 0 to 2π. The distance we have moved on the u-axis after one full rotation, for instance, is just the perimeter $2\pi r$ according to this equation. In fact, since r is kept constant, the u-coordinate is actually just dependent on ϕ: $u(\phi)$.

Based on this information, let us now rewrite the Schrödinger equation in Eq. (4.76) as a function of ϕ rather than u. According to the simple *chain rule*, if $|\psi(u)\rangle$ depends on u, which in turn depends on ϕ, that is $|\psi\rangle = |\psi(u(\phi))\rangle$, then the rate of change of $|\psi\rangle$ with respect to ϕ can be computed as the rate of change of $|\psi\rangle$ with respect to u multiplied by the rate of change of u with respect to ϕ:

$$\frac{d|\psi\rangle}{d\phi} = \frac{d|\psi\rangle}{du}\frac{du}{d\phi}. \tag{4.78}$$

According to Eq. (4.77),

$$\frac{du}{d\phi} = \frac{d}{d\phi}(\phi r) = r. \tag{4.79}$$

Substituting this in Eq. (4.78) gives

$$\frac{d}{d\phi}|\psi\rangle = r\frac{d}{du}|\psi\rangle, \tag{4.80}$$

or simply

$$\frac{d}{d\phi} = r\frac{d}{du}. \tag{4.81}$$

We can now multiply Eq. (4.80) by Eq. (4.81):

$$\begin{aligned}\frac{d}{d\phi}\left(\frac{d}{d\phi}|\psi\rangle\right) &= r\frac{d}{du}\left(r\frac{d}{du}|\psi\rangle\right)\\ \frac{d^2|\psi\rangle}{d\phi^2} &= r^2\frac{d^2|\psi\rangle}{du^2}.\end{aligned} \tag{4.82}$$

This enables us to rewrite the Schrödinger equation of Eq. (4.76) as a function of ϕ:

$$\hat{\mathscr{H}}|\psi(\phi)\rangle = -\frac{\hbar^2}{2mr^2}\frac{d^2|\psi(\phi)\rangle}{d\phi^2} = E|\psi(\phi)\rangle. \tag{4.83}$$

The Hamiltonian is seen to be directly proportional to the square of the angular momentum operator given in Eq. (4.67):

$$\hat{\mathscr{H}} = \frac{1}{2mr^2}\left(-i\hbar\frac{d}{d\phi}\right)^2 = \frac{1}{2mr^2}\hat{L}_z^2. \tag{4.84}$$

The eigenenergies can thus be determined directly by applying the angular momentum operator twice to the eigenfunctions $|\psi_{m_l}\rangle$, or simply $|m_l\rangle$:

$$\begin{aligned}\hat{L}_z^2|m_l\rangle &= \hat{L}_z\left(\hat{L}_z|m_l\rangle\right)\\ &= \hat{L}_z m_l\hbar|m_l\rangle\\ &= m_l\hbar\hat{L}_z|m_l\rangle\\ &= m_l^2\hbar^2|m_l\rangle.\end{aligned} \tag{4.85}$$

One thus obtains for the energy of a particle on a ring:

$$E = \frac{1}{2mr^2}m_l^2\hbar^2 = \frac{m_l^2\hbar^2}{2I}, \tag{4.86}$$

where $I = mr^2$ is the *moment of inertia* of the orbiting particle. Note that states with opposite values of angular momenta have the same energies. They are *degenerate*. The symmetry responsible for this degeneracy may be attributed to the reflection plane, which changes the sign of ϕ; hence,

$$\sigma_v |m_l\rangle = |-m_l\rangle. \tag{4.87}$$

4.5.2 The shell perspective

The solutions of the particle on a ring model, with an almost trivial Lie group at its basis, form an energy spectrum that consists of *shells*. A shell is a universal concept encountered at all quantum levels of matter. It is directly related to the presence of Lie groups and it plays a key role in our treatment. In the current model, the shells are of a straightforward simplicity, and yet they already illustrate nearly all the basic ingredients of the shell concept.

For the particle on a ring, the energy spectrum is as follows: the lowest level, or ground level, corresponds to $m_l = 0$ and is nondegenerate whereas all other levels are twofold degenerate, reflecting the equivalence of clockwise and counterclockwise waves (Figure 4.6). Three ingredients of the symmetry determine the shell structure:

1. The Hamiltonian $\hat{\mathscr{H}}$ characterizes the energy level (i.e., all states in the same energy level are characterized by the same value of m_l^2).
2. The Lie group generator, *in casu* the angular momentum operator $\hat{L}_z$, characterizes the individual *components* of the levels in a unique way. States may be labeled unambiguously by the signed quantum number m_l.
3. Additional symmetry operators, such as σ_v, are available to shift from one component to the others within the same level. Such operators are called *shift* or *ladder operators.*

4.5.3 Aromatic molecules

The Hückel rule

In organic chemistry, *Hückel's rule* determines whether a planar ring molecule has aromatic properties.

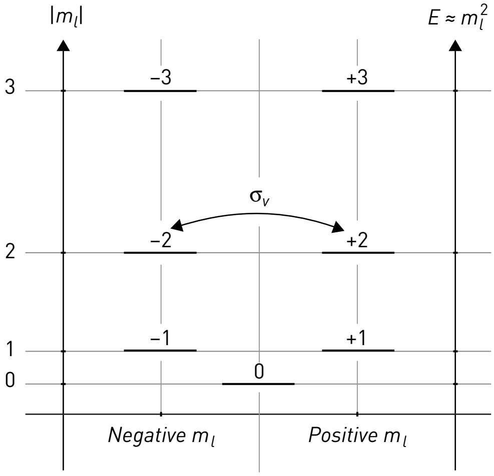


FIGURE 4.6 Energy diagram for a particle on a ring. The first seven energy levels are shown for $m_l = 0, \pm 1, \pm 2, \pm 3$.

Definition 4.2 (Hückel's rule): A cyclic ring molecule is aromatic when the number of its π-electrons equals $4n+2$, where n can be zero or any other positive integer. ▪

For example, the exceptional stability of benzene (the famous Kekulé structure with a ring of six carbon atoms) can be explained by Hückel's rule, which confirms that benzene is an aromatic molecule because of its six π-electrons (for $n=1$, $4n+2=6$). The quantum mechanical basis for its formulation was first worked out by physical chemist Erich Hückel in 1931, based on calculations using the so-called *Hückel method*, but it finds its origin in the particle on a ring model.[12] After all, aromatic molecules are ringlike systems in which the excess valence electrons spin around in both directions (Figure 4.5). The underlying rationale for the number sequence in the Hückel rule is that stable molecules are characterized preferentially by closed-shell electronic structures. This means there should be a large energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). This leads to a sequence that corresponds to the shells of cylindrical quantum systems (described in §4.5.2). Recall from the previous sections that, except for the case $m_l = 0$, there are two quantum states for every value of m_l. The number of energy levels in a shell is thus given by $2n+1$. To fill up all these levels to n then requires $2 \times (2n+1) = 4n+2$ electrons, because electrons can have two additional possible orientations of their spins. The magic numbers of this sequence: 2, 6, 10, 14, 18, and so on thus correspond precisely to the closing of consecutive levels of the cylindrical shell with electron pairs.

When this is clear, a whole family of aromatic compounds fits into a simple numerical scheme. Examples of molecules from 6π to 26π are presented in Table 4.1. Here benzene (6π) and porphyrin (18π) are natural products. Benzene is a constituent of crude oil, and a variety of chemical modifications of porphyrin are found in plants, where they form the chromophoric part of the light-harvesting system. A porphyrin is a small wonder of molecular design shaped by the mechanisms of evolution. At first sight, it bears little resemblance to a chain with eighteen beads, but on closer inspection it is indeed found to incorporate a closed circuit over eighteen atoms. There is a good reason why this circuit is warped and contains bypasses.[13] The cyclic electronic structures require a planar shape, where the $2p_z$-orbitals on neighboring atoms can overlap to enable the hopping of the mobile electrons. But, a straightforward ring of eighteen CH beads, without all the colorful features of the porphyrin, would not be able to sustain the planar geometry and would collapse into a more densely packed cluster structure. Moreover, to harvest the sunlight that penetrates to the surface of the planet, nature had to develop a molecule with a HOMO–LUMO gap in the visible region of the spectrum. According to the ring

---

[12] E. Hückel. "Quantentheoretische Beiträge zum Benzolproblem." *Zeitschrift für Physik* 70 (1931), pp. 204–286; L. Salem. *The Molecular Orbital Theory of Conjugated Systems*. New York: W. A. Benjamin, 1966.

[13] Strictly speaking, Hückel theory is applicable only to monocyclic conjugated systems. However, many polycyclic systems also show aromaticity as a result of internal $4n+2$ circuits. See M. Randić. "Aromaticity of Polycyclic Conjugated Hydrocarbons." *Chemical Reviews* 103 (2003), pp. 3449–3605.

model, the energy of excitation is simply given by the difference between the HOMO and LUMO shells. From Eq. (4.86), we can derive

$$\Delta E = E_{\text{LUMO}} - E_{\text{HOMO}} = \frac{(|m_l| + 1)^2\,\hbar^2}{2I} - \frac{m_l^2\,\hbar^2}{2I}. \tag{4.88}$$

For benzene, the gap is calculated to be in the far-ultraviolet region. As a result, benzene is colorless and cannot act as a chromophore. When going from benzene to porphyrin, two aspects of the ring model change: (1) the size of the ring increases by a factor of three and (2) the excitations are displaced in the shell model. In benzene, the HOMO → LUMO excitation is from $m_l = \pm 1$ to $m_l = \pm 2$; in porphyrin, it is from $m_l = \pm 4$ to $m_l = \pm 5$. In the previous equation, the impact of these changes is an increase of the enumerator by a factor of three as a result of the changes in m_l quantum numbers; on the other hand, however, an increase of the denominator by a factor of nine is seen as a result of the threefold increase of the radius. The net result is a threefold decrease of the gap, shifting the absorption spectrum right into the visible region and endowing porphyrin with the magnificent colors of the leaves.

The other entries in Table 4.1 are synthetic molecules reflecting the art of organic chemistry to design ever-new compounds that incorporate desired properties. The boron subporphyrin is a 14π system.[14] If we wished to synthesize a porphyrin analogue with only three pyrrole rings, but, knowing three is not a divisor of fourteen, the closest we can come to realizing this goal is a subporphyrin with a macrocycle of only twelve atoms. By donating two electrons to the ring, the boron atom in the center stabilizes the system as a dianion, realizing a stable, 14π aromatic ring system.

The 26π system is a hexaphyrin.[15] This is one of the largest synthetic annulene systems thus far. It, too, is aromatic with cyclic delocalization. This molecule exhibits a very interesting feature, which has a peculiar group theoretical significance. When electrochemically reduced by injecting two more electrons, a 28π-electron system is formed, which defies the Hückel rule. Such a system is called *antiaromatic*. Annulenes of this kind carry $4n\pi$-electrons, which implies that the outer shell is only half filled. As a result, such systems are unstable; they are highly reactive and exhibit all kinds of structural distortions. Reduced hexaphyrin is large and flexible enough to distort to a new conformation that is used to explore a different mathematical solution of cyclic symmetry: it spontaneously adopts the twisted conformation of a Möbius ring (Figure 4.7 and Figure 4.8).[16]

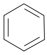
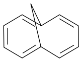
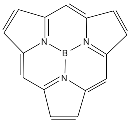
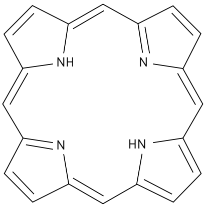
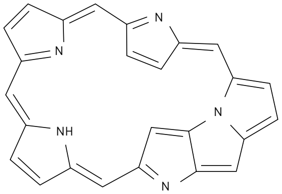
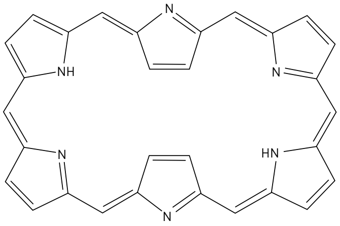
---

[14] T. Torres. "From Subphthalocyanines to Subporphyrins." *Angewandte Chemie, International Edition* 45 (2006), pp. 2834–2837; S. Vancoillie, M. Hendrickx, M. T. Nguyen, K. Pierloot, A. Ceulemans, J. Mack, and N. Kobayashi. "Fourteen-Electron Ring Model and the Anomalous Magnetic Circular Dichroism of Meso-Triarylsubporphyrins." *Journal of Physical Chemistry* A116 (2012), pp. 3960–3967.

[15] A. Jasat and D. Dolphin. "Expanded Porphyrins and their Hetercycles." *Chemical Reviews* 97 (1997), pp. 2267–2340.

[16] J. Sankar, S. Mori, S. Sato, H. Rath, M. Suzuki, Y. Inokuma, H. Shinokubo, K. S. Kim, Z. S. Yoon, J.-Y. Shin, J. M. Lim, Y. Matsuzaki, O. Matsushita, A. Muranaka, N. Kobayashi, D. Kim, and A. Osuka. "Unambiguous Identification of Möbius Aromaticity for Meso-Aryl-Substituted [28] Hexaphyrins." *Journal of the American Chemical Society* 130 (2008), pp. 13568–13579.

Table 4.1 Aromatic molecules with 6π- to 26π-electrons

n	$4n+2$	*Name*	*Molecular structure*
1	6	Benzene	
2	10	1,6-methano[10]annulene	
3	14	Boron subporphyrin	
4	18	Porphyrin	
5	22	*N*-fused[22]pentaphyrin	
6	26	Hexaphyrin	

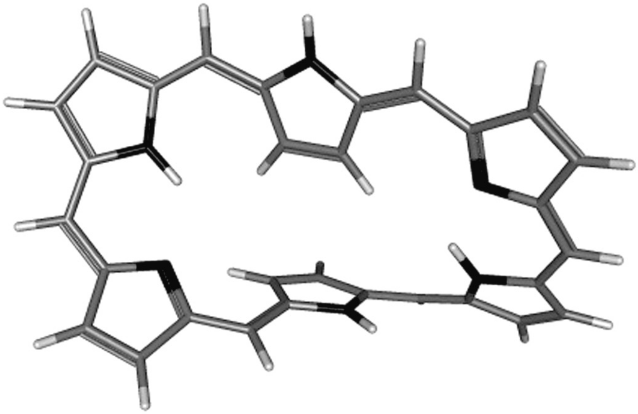


FIGURE 4.7 Structure of an electrochemically reduced hexaphyrin molecule. As a result of the presence of 28π-electrons, the molecule spontaneously adopts a twisted Möbius—like conformation.

FIGURE 4.8 Escher's *Möbius Strip II* is a woodcut from 1963. The *Möbius strip* (variously known as a *Möbius band* or *twisted cylinder*) was first discovered by the German mathematician August Ferdinand Möbius (1790–1868) in 1858. You can create your own Möbius band easily by cutting out a single strip of paper, and giving one of its ends a half twist before gluing the two ends to create a loop. What is intriguing about a *Möbius band* is that it is a one-sided surface, with only one boundary! To convince yourself of this fact, try following one of the nine ants in the figure on their journey along the band. By crawling one full length of the original strip, each ant ends up on "the other side" of where it started. By covering the same length once again, it ultimately returns to its starting position. Escher thus noted: "An endless ring-shaped band usually has two distinct surfaces, one inside and one outside. Yet on this strip nine ants crawl after each other and travel the front side as well as the reverse side. Therefore the strip has only one surface."[17] Notice that with a cylinder, the ants would have to *cross an edge* to reach the other side. Funnily enough, by placing the band on its side, one obtains an inverted number 8 or *lemniscate* ∞, the symbol for infinity. Reprinted with permission of the Escher Foundation, Den Haag.

The Heilbronner rule

As is well known, the topology of a Möbius band requires that the original starting configuration is restored only after revolving twice around the center. As a result, a wave function depending on the angle of revolution ϕ is periodic in $\phi + 4\pi$. Hence, Eq. (4.70) has to be recast as follows:

$$\begin{aligned} |\psi(\phi)\rangle &= |\psi(\phi + 4\pi)\rangle \\ e^{ik\phi} &= e^{ik\phi}\, e^{ik4\pi} \\ 1 &= e^{ik4\pi}. \end{aligned} \tag{4.89}$$

This equation is now satisfied for half integer values of k. The magnetic quantum numbers for cyclic waves on a Möbius band thus are

$$m_l = \pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{5}{2}, \ldots. \tag{4.90}$$

The resulting shell structure is thus at odds with the Hückel rule, predicting closed shells for systems with $4n$ electrons rather than $2 + 4n$. The electron count for Möbius

---

[17] M. C. Escher. *The Graphic Work*. Köln: Taschen, 2001, p. 12.

annulenes was predicted by Heilbronner as early as 1964,[18] but it took about forty years before the first molecules with this topology were synthesized.

Definition 4.3 (Heilbronner rule): An annulene molecule with a Möbius band topology is aromatic when the number of its π-electrons equals $4n$, where n is any positive integer. ■

The remarkable property of the reduced hexaphyrin is that it spontaneously seeks the twisted geometry to fulfil the Heilbronner rule. In smaller rings this doesn't work because the twist introduces too much steric tension.

The half-integer quantum solutions of the rotational symmetry group are part of the theory of angular momentum. They have been waiting patiently "in potentiality" as mathematical solutions for more than a century, until they finally reached their "actuality" in a physical system such as the hexaphyrin molecule. They will make their official appearance in the next chapter, where the rotation group will be extended to rotations in three dimensions.

---

[18] E. Heilbronner. "Hückel Molecular Orbitals of Möbius-Type Conformations of Annulenes." *Tetrahedron Letters* 5.29(1964), pp. 1923–1928.

5 The SO(3) group

Eppur si muove.[1]
–Galileo Galilei (1563–1642)

At the foundations of modern cosmology (and the Big Bang theory in particular) lies the *Cosmological Principle*, which assumes that space is *homogeneous* and *isotropic* on sufficiently large scales. The homogeneity of space means the universe looks and behaves the same from every possible location; all locations are said to be *equivalent* in the sense that our universe does not possess any priviledged position or universal center. The isotropy of space, on the other hand, means that wherever you look (up or down, left or right), the universe always looks the same, implying there is no preferred direction in space, and that all directions are therefore completely *equivalent.*

The homogeneity of space implies the universe has *translational invariance* (i.e., the laws of physics are invariant under spatial translations). The isotropy of space implies the universe has *rotational invariance* (i.e., the laws of physics are invariant under rotational transformations). Therefore, physical experiments have the same outcome regardless of their location and orientation in space.

On a microscopic scale, the uniformity and isotropy of space implies that the Hamiltonian of an isolated classical system is *invariant* under three-dimensional rotations; that is, the mathematical form of the Hamiltonian (and therefore also the

---

[1] "And yet it moves."

energy of the system) does not change under spatial rotations of the system. One says that the Hamiltonian has *spherical symmetry*.[2]

Interestingly, the spherical symmetry properties of the Hamiltonian are responsible for the *conservation of the total angular momentum* **L** of the system. This is an example of *Noether's theorem*, which states that for every continuous symmetry of a system, there is a corresponding conservation law, and vice versa. This idea, which first originated in the works of Emmy Noether, has proved to be of such fundamental importance in modern physics that it will often be at the center of our current discussion as well. Not surprisingly, this chapter intends to fully explore the intimate relationship that exists between three-dimensional rotations on the one hand, and the angular momentum of a physical system on the other. In the subsequent scholium, we will pause to contemplate Noether's theorem in more detail.

Chapter outline

The chapter proceeds along the same logic as was for the preceding presentation of the SO(2) group. We start our discussion in §5.1 by defining the symmetry group of the sphere from a geometric point of view as the group of three-dimensional rotations of the sphere. We then introduce the matrices, corresponding to these rotations, and show that they are orthogonal. We continue our exploration in §5.2 with the introduction of the algebraic group of 3×3 orthogonal matrices, labeled O(3). It turns out that the matrices describing rotations of the sphere coincide with the elements of the special orthogonal subgroup SO(3) of orthogonal 3×3 matrices with determinant $+1$. In §5.3 we derive the general expression for the rotation matrices using a special property of orthogonal matrices due to Jacobi.

Having thus established the central role of SO(3) in the treatment of spherical systems, we are then ready to focus on the Lie group characteristics of SO(3) (with its elements being characterized by a set of three continuously varying parameters) and its corresponding Lie algebra. In the cyclic problem of Chapter 4, we already saw a glimpse of the power of a Lie algebra, albeit the $\mathfrak{so}(2)$ algebra was almost trivial because it contained only one infinitesimal operator. The spherical symmetry group SO(3) exhibits new features because it is generated by three generators, as we will see in § 5.4.

The fundamental link between the group theoretical generators of the $\mathfrak{so}(3)$ algebra and the three Cartesian components of the angular momentum operator $\hat{L}$ are brought to light in §5.5 when we examine rotations in quantum mechanics. In fact, the resulting angular momentum algebra is sufficient to obtain all the allowed eigenstates and eigenvalues of the angular momentum observables, as will be outlined in §5.6. In §5.7, we derive the Hamiltonian of a free particle on a sphere and examine its spherical ingredients. The results are compared with the electronic structure of the celebrated carbon buckyball.

5.1 THE SPHERICAL ROTATION GROUP

Let $\hat{R}$ denote a physical rotation in three-dimensional Euclidean space $\mathbb{R}^3$. The set of all rotation operators $\{\hat{R}\}$ forms an infinite continuous Lie group—the *spherical rotation group*. This can be verified by checking the four group axioms:

---

[2] This is, of course, no longer the case when a preferred direction in space is imposed by applying an external (electric, magnetic, or gravitational) field.

1. *Closure*: To start with, we note that for every two rotations $\hat{R}_1$ and $\hat{R}_2$, there is a third rotation $\hat{R}_3$, such that

$$\hat{R}_3 = \hat{R}_2 \star \hat{R}_1, \tag{5.1}$$

where the product symbol $\star$ is used to denote the consecutive application of $\hat{R}_1$ and $\hat{R}_2$. Although intuitively obvious—after all, no matter how often you rotate a sphere, the result is always another rotation—the first mathematical proof of Eq. (5.1) only appeared in 1775 in a paper by Leonhard Euler (1707–1783). Later, Olinde Rodrigues offered an equivalent proof, but whereas Euler had used an algebraic approach, Rodrigues proceeded by pure geometric construction.[3]

2. *Associativity*: For all $\hat{R}_1$, $\hat{R}_2$, and $\hat{R}_3$, the associative law holds true:

$$\left(\hat{R}_3 \star \hat{R}_2\right) \star \hat{R}_1 = \hat{R}_3 \star \hat{R}_2 \star \hat{R}_1 = \hat{R}_3 \star (\hat{R}_2 \star \hat{R}_1). \tag{5.2}$$

3. *Identity element*: There exists an element $\hat{E}$ such that for every element $\hat{R}$,

$$\hat{E} \star \hat{R} = \hat{R} \star \hat{E} = \hat{R}. \tag{5.3}$$

Clearly, the identity element $\hat{E}$ corresponds to the trivial rotation through 0°.

4. *Inverse element*: Finally, every rotation $\hat{R}$ can be nullified by a rotation $\hat{R}'$ in the opposite sense; that is, $\hat{R}'$ is the inverse element $\hat{R}^{-1}$ of $\hat{R}$:

$$\hat{R} \star \hat{R}^{-1} = \hat{R}^{-1} \star \hat{R} = \hat{E}. \tag{5.4}$$

This completes our proof that the set of three-dimensional rotations forms a group—the *spherical rotation group*. To assign a rotation in three dimensions, we must specify the axis and angle of rotation. Every rotation is effected around a certain *rotation axis* that passes through the origin O, and has a direction that can be specified by the vector $\boldsymbol{n}$. This vector is determined by the directional cosines n_x, n_y, and n_z:

$$\boldsymbol{n} = n_x \boldsymbol{e}_x + n_y \boldsymbol{e}_y + n_z \boldsymbol{e}_z. \tag{5.5}$$

Moreover, the vector $\boldsymbol{n}$ is normalized:

$$|\boldsymbol{n}| = \sqrt{n_x^2 + n_y^2 + n_z^2} = 1. \tag{5.6}$$

The scalar components n_x, n_y, and n_z of $\boldsymbol{n}$ may be expressed by the two angle coordinates, θ and ϕ, which describe the latitude and longitude, respectively:

$$\begin{aligned} n_x &= \sin\theta \cos\phi; \\ n_y &= \sin\theta \sin\phi; \\ n_z &= \cos\theta. \end{aligned} \tag{5.7}$$

So in reality, only two parameters (θ and ϕ) are needed to specify the orientation of the rotation axis $\boldsymbol{n}$. Finally, we need to specify the angle over which the sphere is rotated. This *rotation angle* is denoted as ω. As a result, *only three effective parameters* are needed to specify a rotation in three dimensions. The corresponding operation is

---

[3] For a detailed description of this intriguing history, see S. L. Altmann. *Rotations, Quaternions, and Double Groups*. Oxford: Clarendon Press, 1986.

parametrized as $\hat{R}(\omega\boldsymbol{n})$. This manner of parameterizing three-dimensional rotations is known as the *axis-angle parameterization.* Notice also that the sense of the rotation (i.e., clockwise or counterclockwise) can always be determined with the help of the right-hand rule (that is, if you wrap your right hand around the rotation axis so that your thumb is pointing in the direction of $\boldsymbol{n}$, then your fingers indicate the direction of a positive rotation angle).

It should also be mentioned that the mapping between the group of physical rotations and the axis-angle parameter set is not one to one. A given rotation can be specified by four different combinations of parameters:

$$\begin{array}{ll} 1. & \hat{R}(\omega, n_x, n_y, n_z) \\ 2. & \hat{R}(-\omega, -n_x, -n_y, -n_z) \\ 3. & \hat{R}(-2\pi + \omega, n_x, n_y, n_z) \\ 4. & \hat{R}(2\pi - \omega, -n_x, -n_y, -n_z) \end{array} \tag{5.8}$$

Combinations 1 and 2 give rise to the same rotation because both the angle and the direction change sign simultaneously, as do combinations 3 and 4. Physically, they are identical viewed only from opposite poles. A counterclockwise rotation viewed from the north pole is viewed as clockwise from the south pole, but it is the same rotation. On the other hand, the physical rotation for 1 (and 2) differs from the one described by 3 (and 4), despite having the same outcome. In 1 the rotation is counterclockwise over angle ω, whereas in 3 one rotates about the same rotation axis but in the opposite direction over an angle $2\pi - \omega$ to arrive at the same point. The mapping between points is thus the same, but the *path* that is followed is different. This aspect is not taken into account in this treatment of SO(3), but may nevertheless be important when transporting quantum states.

5.2 THE ORTHOGONAL GROUP IN THREE DIMENSIONS

So far, we have viewed rotations as actual motions of the sphere. We now consider the effect of rotations on the position vectors of points on a sphere and express this by means of rotation matrices. Rotations should leave the scalar products of the position vectors invariant, implying that the rotation matrices should be orthogonal matrices. In this way, we arrive at the group of all three-dimensional orthogonal matrices, denoted as the orthogonal group in three dimensions: O(3).

5.2.1 Rotation matrices

Let $\mathbf{a}$ represent a position vector in real three-dimensional Euclidean space $\mathbb{R}^3$ with a tail that coincides with the origin O of a Cartesian $\Sigma\left(x, y, z\right)$ reference system, and with a head that points to point P, defined by the set of Cartesian coordinates $\left(a_x, a_y, a_z\right)$. In matrix notation,

$$\mathbf{a} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}. \tag{5.9}$$

In general, all rotations are defined relative to a center of rotation, which is usually chosen to be the origin O, which remains fixed during all rotations. When $\mathbf{a}$ is spatially rotated about the origin O, a new vector $\mathbf{a}'$ is obtained with Cartesian coordinates $\left(a'_x, a'_y, a'_z\right)$:

$$\mathbf{a}' = \begin{bmatrix} a'_x \\ a'_y \\ a'_z \end{bmatrix}. \tag{5.10}$$

This rotation is induced by a rotation operator, which we denote with the help of the axis-angle parameters as $R(\omega\boldsymbol{n})$.

$$\mathbf{a}' = R(\omega\boldsymbol{n})\,\mathbf{a}. \tag{5.11}$$

The operator $R(\omega\boldsymbol{n})$ induces a linear transformation[4] of the coordinates and can therefore be expressed by a 3×3 *rotation matrix* $\mathbb{R}(\omega\boldsymbol{n})$:

$$\mathbf{a}' = \mathbb{R}(\omega\boldsymbol{n})\,\mathbf{a}. \tag{5.12}$$

The set of rotation matrices $\{\mathbb{R}(\omega\boldsymbol{n})\}$ forms a *matrix representation* of the spherical rotation group. Recall also from Chapter 3, §3.3, that when a rotation matrix acts on coordinates, it always precedes the column vector of coordinates; when acting on coordinate functions, it follows the row vector of these functions.

When two vectors **a** and **b** are rotated, their *length* as well as the *angle* γ between them remains invariant. Therefore, rotations conserve the *scalar product*:

$$\mathbf{a}' \cdot \mathbf{b}' = \left|\mathbf{a}'\right| \left|\mathbf{b}'\right| \cos\gamma = |\mathbf{a}|\,|\mathbf{b}| \cos\gamma = \mathbf{a} \cdot \mathbf{b}. \tag{5.13}$$

In matrix notation,

$$\mathbf{a}' \cdot \mathbf{b}' = \mathbf{a}'^{\mathrm{T}}\mathbf{b}' = \mathbf{a}^{\mathrm{T}}\mathbf{b} = \mathbf{a} \cdot \mathbf{b}. \tag{5.14}$$

Since $\mathbf{a}' = \mathbb{R}(\omega\boldsymbol{n})\,\mathbf{a}$ and $\mathbf{b}' = \mathbb{R}(\omega\boldsymbol{n})\,\mathbf{b}$, we have (leaving out the axis-angle parameterization for notational simplicity):

$$\mathbf{a}'^{\mathrm{T}}\mathbf{b}' = [\mathbb{R}\mathbf{a}]^{\mathrm{T}}\,\mathbb{R}\mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbb{R}^{\mathrm{T}}\mathbb{R}\mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b}. \tag{5.15}$$

It follows from the invariance of the scalar product that

$$\mathbb{R}^{\mathrm{T}}\mathbb{R} = \mathbb{I}. \tag{5.16}$$

The order in this product can also be inverted. To this aim, we multiply to the left by $\mathbb{R}$ and to the right by $\mathbb{R}^{\mathrm{T}}$:

$$\begin{gathered} \mathbb{R}\mathbb{R}^{\mathrm{T}}\mathbb{R}\mathbb{R}^{\mathrm{T}} = \mathbb{R}\mathbb{R}^{\mathrm{T}} \\ \mathbb{R}\mathbb{R}^{\mathrm{T}}\mathbb{R}\mathbb{R}^{\mathrm{T}}\left[\mathbb{R}\mathbb{R}^{\mathrm{T}}\right]^{-1} = \mathbb{R}\mathbb{R}^{\mathrm{T}}\left[\mathbb{R}\mathbb{R}^{\mathrm{T}}\right]^{-1} \\ \mathbb{R}\mathbb{R}^{\mathrm{T}} = \mathbb{I}. \end{gathered} \tag{5.17}$$

The last step in this proof is based on the assumption that $\mathbb{R}\mathbb{R}^{\mathrm{T}}$ is nonsingular, so that the inverse of this matrix product exists. This is verified easily from Eq. (5.16) by taking the determinants. Since the unit matrix $\mathbb{I}$ has unit determinant (i.e., $\det\mathbb{I} = 1$), it follows from Eq. (5.16) that

$$\det\left[\mathbb{R}^{\mathrm{T}}\mathbb{R}\right] = \det\mathbb{R}^{\mathrm{T}}\det\mathbb{R} = [\det\mathbb{R}]^2 = 1, \tag{5.18}$$

---

[4] The transformation is said to be *linear* because the rotation of a sum of vectors is equal to the sum of the rotated vectors:

$$R(\omega\boldsymbol{n})\,(\mathbf{a} + \mathbf{b}) = R(\omega\boldsymbol{n})\,\mathbf{a} + R(\omega\boldsymbol{n})\,\mathbf{b} = \mathbf{a}' + \mathbf{b}'.$$

where use is made of the fact that $\det \mathbb{A}\mathbb{B} = \det \mathbb{A} \det \mathbb{B}$, and that $\det \mathbb{A}^{\mathrm{T}} = \det \mathbb{A}$. Hence, the determinant of the $\mathbb{R}$ matrix equals ±1. Matrices satisfying both the conditions in Eqs. (5.16) and (5.17) are said to be *orthogonal*. Rows and columns in such matrices are orthonormal to each other, meaning that the scalar product of two rows or two columns is equal to zero, except for the products of a row with itself or a column with itself.

5.2.2 The orthogonal group O(3)

Definition 5.1 (The O(3) Lie group): The set of all 3 × 3 orthogonal matrices:

$$\mathrm{O}(3) = \left\{ \text{real } 3 \times 3 \text{ matrices } \mathbb{A} : \begin{array}{l} \mathbb{A}^{\mathrm{T}}\mathbb{A} = \mathbb{A}\mathbb{A}^{\mathrm{T}} = \mathbb{I} \\ |\det \mathbb{A}| = 1 \end{array} \right\} \tag{5.19}$$

forms an infinite Lie group under matrix multiplication. This group is called the *orthogonal group* in three dimensions, and is denoted by the symbol O(3). ■

Proof. To qualify as a group, O(3) should satisfy the four *group axioms*, as defined in Chapter 2. This can be easily verified by letting $\mathbb{A}_a, \mathbb{A}_b \in \mathrm{O}(3)$ denote two orthogonal matrices, and assuming that $\mathbb{A}_c = \mathbb{A}_a\mathbb{A}_b$.

1. *Closure*: One needs to prove that $\mathbb{A}_c \in \mathrm{O}(3)$; in other words, $\mathbb{A}_c$ is a real 3 × 3 orthogonal matrix with a unimodular determinant. Since $\mathbb{A}_a$ and $\mathbb{A}_b$ are real 3 × 3 matrices, $\mathbb{A}_c$ is also a real 3 × 3 matrix. Also,

$$\mathbb{A}_c^{\mathrm{T}}\mathbb{A}_c = (\mathbb{A}_a\mathbb{A}_b)^{\mathrm{T}}\mathbb{A}_a\mathbb{A}_b = \mathbb{A}_b^{\mathrm{T}}\mathbb{A}_a^{\mathrm{T}}\mathbb{A}_a\mathbb{A}_b = \mathbb{A}_b^{\mathrm{T}}\mathbb{A}_b = \mathbb{I} \text{ (orthogonality)}. \tag{5.20}$$

2. *Associativity*: Matrix multiplication is associative, so the associative law obviously holds for the O(3) group elements.
3. *Identity element*: The 3 × 3 identity matrix $\mathbb{I}$ represents the identity element.
4. *Inverse element*: Let $\mathbb{A}_a^{-1}$ denote the inverse matrix of $\mathbb{A}_a$. We need to prove that $\mathbb{A}_a^{-1} \in \mathrm{O}(3)$. Since $(\mathbb{A}_a^{-1})^{\mathrm{T}} = (\mathbb{A}_a^{\mathrm{T}})^{-1}$,

$$\left(\mathbb{A}_a^{-1}\right)^{\mathrm{T}}\mathbb{A}_a^{-1} = \left(\mathbb{A}_a^{\mathrm{T}}\right)^{-1}\mathbb{A}_a^{-1} = \left(\mathbb{A}_a\mathbb{A}_a^{\mathrm{T}}\right)^{-1} = \mathbb{I}^{-1} = \mathbb{I} \text{ (orthogonality)}. \tag{5.21}$$

This proves that O(3) forms a group. ■

5.2.3 The special orthogonal group SO(3)

The group O(3) is disconnected because it consists of two kind of elements: (1) orthogonal matrices with determinant +1 and (2) orthogonal matrices with determinant −1. The former set forms a subgroup of O(3), because the product of two matrices of this set always yields a resulting matrix with determinant +1. This subgroup is known as the *special orthogonal group*, or SO(3). The set of matrices with determinant −1 is known in group theory as the *coset* of this group. As we shall see, the SO(3) group in fact corresponds to the group of proper rotations whereas its coset collects all improper rotations, covering reflections, rotation-reflections, and inversion.

5.3 ROTATIONS AND SO(3)

It is now time to confront the rotation group with the group of orthogonal matrices by deriving explicitly the matrices, $\mathbb{R}(\omega \boldsymbol{n})$, that represent the actual rotation operations. This derivation is more involved than the case of rotations in two dimensions, which depends on one parameter only. The derivation not only aims at the desired matrix forms, but also demonstrates that *every orthogonal matrix with determinant* $+1$ *is a rotation matrix.*

5.3.1 Orthogonality and skew-symmetry

To start, we note that an orthogonal matrix $\mathbb{A}$ can be written as the exponential of a 3×3 matrix $\mathbb{S}$:

$$\mathbb{A} = \exp \mathbb{S} \tag{5.22}$$

By explicit expansion of the exponential, we see that the transpose relation also holds:

$$\mathbb{A}^{\mathrm{T}} = \exp\left(\mathbb{S}^{\mathrm{T}}\right). \tag{5.23}$$

Since $\mathbb{A}$ is orthogonal, we obtain

$$\mathbb{A}\mathbb{A}^{\mathrm{T}} = \mathbb{I} = \exp \mathbb{S} \, \exp\left(\mathbb{S}^{\mathrm{T}}\right) = \exp\left(\mathbb{S} + \mathbb{S}^{\mathrm{T}}\right). \tag{5.24}$$

This equation can be fulfilled only if the sum on the right-hand side of this equation is the null matrix, $\mathbb{O}$. Thus, we have

$$\begin{aligned} \mathbb{O} &= \mathbb{S} + \mathbb{S}^{\mathrm{T}} \\ \mathbb{S} &= -\mathbb{S}^{\mathrm{T}}. \end{aligned} \tag{5.25}$$

A matrix that changes its sign on transposition is called a *skew-symmetric* (or *antisymmetric*) matrix. In terms of its matrix elements, $S_{ij} = -S_{ji}$. For the diagonal elements of $\mathbb{S}$, this property leads to $S_{ii} = -S_{ii}$, which holds only for $S_{ii} = 0$. The general form of an $n \times n$ skew-symmetric matrix $\mathbb{S}$ is therefore given by

$$\mathbb{S} = \begin{bmatrix} 0 & S_{12} & S_{13} & \cdots & S_{1n} \\ -S_{12} & 0 & S_{23} & \cdots & S_{2n} \\ -S_{13} & -S_{23} & 0 & \cdots & S_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -S_{1n} & -S_{2n} & -S_{3n} & \cdots & 0 \end{bmatrix}. \tag{5.26}$$

The matrix $\mathbb{S}$ has, in general, n^2 matrix elements. The n diagonal elements $\mathbb{S}_{ii}$ are all zero, leaving $n^2 - n$ unique matrix elements. However, because of the skew-symmetry of $\mathbb{S}$, the lower triangular part of $\mathbb{S}$ is perfectly defined by the upper triangular part. The number N of unique matrix elements of an $n \times n$ skew-symmetric matrix $\mathbb{S}$ is therefore given by the simple formula

$$N = \frac{n^2 - n}{2} = \frac{n(n-1)}{2}. \tag{5.27}$$

As a further consequence, the trace, which is the sum of the diagonal elements, is also zero:

$$\mathrm{Tr}(\mathbb{S}) = \sum_{i=1}^{n} S_{ii} = 0. \tag{5.28}$$

Now, according to a relationship from Jacobi, the determinant of the orthogonal matrix $\mathbb{A}$ is obtained by exponentiating the trace of the skew-symmetric matrix $\mathbb{S}$, which yields

$$\det \mathbb{A} = \exp(\mathrm{Tr}(\mathbb{S})) = e^0 = 1. \tag{5.29}$$

For a derivation of this identity, see Appendix E. Let us summarize what we have so far: any orthogonal matrix with determinant +1 can be written as the exponent of a skew-symmetric matrix. All SO(3) matrices are thus defined as the exponent of a 3 × 3 skew-symmetric matrix; hence,

$$\mathbb{A} = \exp \begin{bmatrix} 0 & S_{12} & S_{13} \\ -S_{12} & 0 & S_{23} \\ -S_{13} & -S_{23} & 0 \end{bmatrix}. \tag{5.30}$$

The parameter space of this matrix is determined by only three real parameters. By varying these numbers independently, we cover the entire set of SO(3) elements. If we now let these parameters become infinitesimally small, the expression can be approximated to first order as

$$\mathbb{A} = \mathbb{I} + \begin{bmatrix} 0 & dS_{12} & dS_{13} \\ -dS_{12} & 0 & dS_{23} \\ -dS_{13} & -dS_{23} & 0 \end{bmatrix}. \tag{5.31}$$

Here, the unit matrix, $\mathbb{I}$, corresponds to the zeroth-order term of the exponential, and the $\mathbb{S}$ matrix itself is the first-order term. This result forms the basis for the construction of the rotation matrices.

5.3.2 The matrix representing an infinitesimal rotation

We now connect the algebraic expression for an orthogonal matrix in the neighborhood of the unit element to an infinitesimal rotation over an angle $d\omega$. At the start, the rotation axis is oriented along the z-axis, with $\boldsymbol{n} = [0,0,1]^T$. Figure 5.1 shows the change of the x-coordinate function as a small vector parallel to the positive y-axis. This change can be written as

$$dx = y\, d\omega. \tag{5.32}$$

We now tilt the pole of the rotation axis in the direction of the y-axis so that the direction becomes $\boldsymbol{n} = [0, n_y, n_z]^T$. As Figure 5.1 indicates, this implies that the rotational plane is tilted around the x-direction. As a result, the displacement of x is tilted downward, too, over the same angle as the pole of the rotation axis itself. This displacement is thus expressed as

$$dx = n_z y\, d\omega - n_y z\, d\omega. \tag{5.33}$$

On the other hand, an inclination of the rotation axis toward the x-direction (i.e., for $n_x \neq 0$) does not introduce an additional contribution to dx. Similar considerations can be made for the other Cartesian directions. It is not difficult to infer that, for an arbitrary orientation of the rotation axis, $\boldsymbol{n} = [n_x, n_y, n_z]^T$, the overall change of the coordinate functions is given by

$$\begin{aligned} dx &= n_z y\, d\omega - n_y z\, d\omega \\ dy &= n_x z\, d\omega - n_z x\, d\omega \\ dz &= n_y x\, d\omega - n_x y\, d\omega. \end{aligned} \tag{5.34}$$

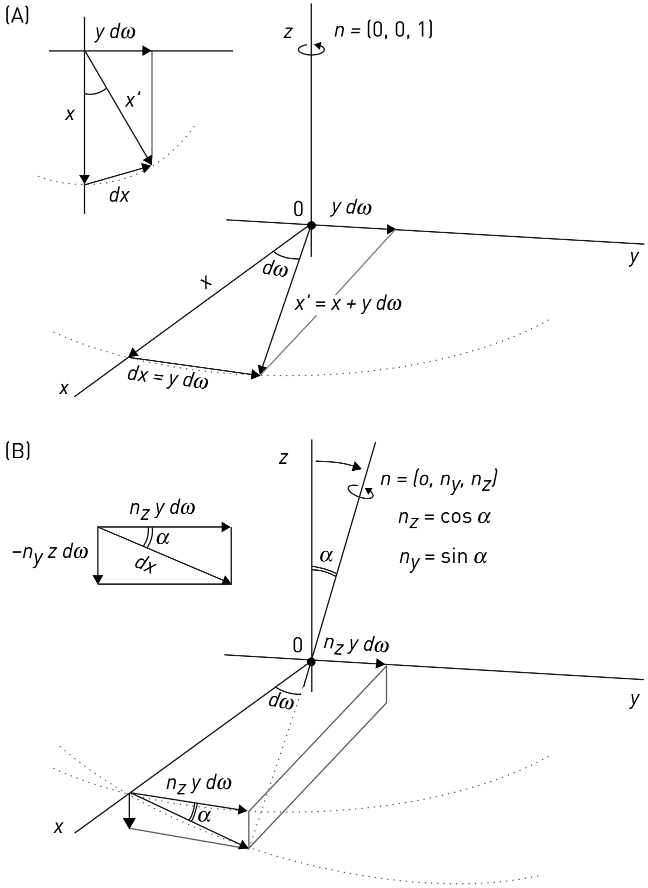


FIGURE 5.1 Infinitesimal rotation of the *x*-coordinate function. (A) In this view, *x* is rotated over $d\omega$ about the *z*-axis and thereby acquires a fraction of *y*-character, equal to $yd\omega$. (B) If the rotation axis is tilted over an angle α toward *y*, the change of the *x* function also gains some *z*-character, equal to $-n_y z d\omega$.

It should be emphasized that this simple linear relationship is made possible by the infinitesimal nature of the changes, whereby all higher order terms can be omitted. This is the essential insight of Lie's treatment of continuous operations. The result can be rewritten as

$$\hat{R}(d\omega\,\boldsymbol{n})\begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} x & y & z \end{bmatrix}\left(\mathbb{I} + d\omega \begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix}\right). \quad (5.35)$$

By comparing this result with Eq. (5.31), the rotation parameters can be mapped unambiguously onto the three independent parameters of an infinitesimal SO(3) matrix. Vice versa, every SO(3) matrix can be identified with a specific rotation matrix:

$$\begin{aligned} dS_{12} &= -d\omega\, n_z \\ dS_{13} &= d\omega\, n_y \\ dS_{23} &= -d\omega\, n_x. \end{aligned} \quad (5.36)$$

The rotation of the coordinate functions may be rewritten in differential form as

$$\frac{d}{d\omega}\hat{R}(d\omega\,\boldsymbol{n})\begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} x & y & z \end{bmatrix}\mathbb{Z}, \tag{5.37}$$

with

$$\mathbb{Z} = \begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix}. \tag{5.38}$$

Here, $\mathbb{Z}$ is the generator matrix of the infinitesimal rotation.

5.3.3 The exponential map

Finally, we extrapolate the infinitesimal operator to a finite rotation, which is provided by the *exponential map*. Recall that for the SO(2) group, the generator was defined in Chapter 4 as the derivative of the 2 × 2 rotation matrix $\mathbb{R}(\omega)$ with respect to ω taken at $\omega = 0$:

$$\mathbb{X} = \left.\frac{d}{d\omega}\mathbb{R}(\omega)\right|_{\omega=0}. \tag{5.39}$$

An arbitrary rotation $\mathbb{R}(\omega)$ could then be "generated" from the exponentiation of the generator $\mathbb{X}$ (a Lie algebra element). That is, if $\mathbb{X} \in \mathfrak{so}(2)$, then

$$\mathbb{R}(\omega) = e^{\omega\mathbb{X}} = \sum_{n=0}^{\infty}\frac{(\omega\mathbb{X})^n}{n!} \in \mathrm{SO}(2). \tag{5.40}$$

Definition 5.2 (SO(2) exponential map): There exists an *exponential map* $\exp : \mathfrak{so}(2) \rightarrow \mathrm{SO}(2)$, given by

$$\mathbb{R}(\omega) = e^{\omega\mathbb{X}}, \tag{5.41}$$

that relates the generator of the $\mathfrak{so}(2)$ *Lie algebra* with the elements of the SO(2) *Lie group* by expressing the rotation matrix $\mathbb{R}(\omega)$ in terms of the infinitesimal generator $\mathbb{X}$. ■

In SO(3), the generator matrix for a rotation in three dimensions was identified in Eq. (5.37) as the $\mathbb{Z}$ matrix. Executing this rotation over a finite angle ω will thus be exponentiated as

$$\mathbb{R}(\omega\boldsymbol{n}) = e^{\omega\mathbb{Z}}. \tag{5.42}$$

This exponential map can be resolved further as a result of a remarkable property of the $\mathbb{Z}$ matrix. Indeed, we can easily verify that

$$\mathbb{Z}^2 = \begin{bmatrix} -n_y^2 - n_z^2 & n_x n_y & n_x n_z \\ n_x n_y & -n_x^2 - n_z^2 & n_y n_z \\ n_x n_z & n_y n_z & -n_x^2 - n_y^2 \end{bmatrix}$$

$$\mathbb{Z}^3 = -(n_x^2 + n_y^2 + n_z^2)\begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix} = -\mathbb{Z}, \tag{5.43}$$

where use was made of Eq. (5.6). The power series for the exponential is then simplified as follows:

$$\begin{aligned}
e^{\omega\mathbb{Z}} &= \mathbb{I}+\omega\mathbb{Z}+\frac{1}{2}\omega^2\mathbb{Z}^2+\frac{1}{3!}\omega^3\mathbb{Z}^3+\frac{1}{4!}\omega^4\mathbb{Z}^4+\frac{1}{5!}\omega^5\mathbb{Z}^5+\frac{1}{6!}\omega^6\mathbb{Z}^6+\cdots \\
&= \mathbb{I}+\omega\mathbb{Z}+\frac{1}{2}\omega^2\mathbb{Z}^2-\frac{1}{3!}\omega^3\mathbb{Z}-\frac{1}{4!}\omega^4\mathbb{Z}^2+\frac{1}{5!}\omega^5\mathbb{Z}+\frac{1}{6!}\omega^6\mathbb{Z}^2-\cdots \\
&= \mathbb{I}+\left(\omega-\frac{1}{3!}\omega^3+\frac{1}{5!}\omega^5-\cdots\right)\mathbb{Z}+\left(\frac{1}{2}\omega^2-\frac{1}{4!}\omega^4+\frac{1}{6!}\omega^6-\cdots\right)\mathbb{Z}^2 \\
&= \mathbb{I}+\sin\omega\,\mathbb{Z}+(1-\cos\omega)\,\mathbb{Z}^2 \\
&= \mathbb{I}+\sin\omega\,\mathbb{Z}+2\sin^2(\omega/2)\,\mathbb{Z}^2.
\end{aligned} \tag{5.44}$$

Here, the power series in odd and even powers is recognized as the series expansion on sin and cos functions, respectively. Finally, the desired general expression of the rotation matrix in three dimensions reads

$$\mathbb{R}(\omega\boldsymbol{n}) =
\begin{pmatrix}
1-2(n_y^2+n_z^2)\sin^2(\omega/2) & -n_z\sin\omega+2n_xn_y\sin^2(\omega/2) & n_y\sin\omega+2n_zn_x\sin^2(\omega/2) \\
n_z\sin\omega+2n_xn_y\sin^2(\omega/2) & 1-2(n_z^2+n_x^2)\sin^2(\omega/2) & -n_x\sin\omega+2n_yn_z\sin^2(\omega/2) \\
-n_y\sin\omega+2n_zn_x\sin^2(\omega/2) & n_x\sin\omega+2n_yn_z\sin^2(\omega/2) & 1-2(n_x^2+n_y^2)\sin^2(\omega/2)
\end{pmatrix}. \tag{5.45}$$

It can be easily verified that this matrix is indeed orthogonal and that its determinant is equal to $+1$. The matrix $\mathbb{Z}$ contains three directional parameters—n_x, n_y, n_z—but these are not independent because their norm is equal to one. However, multiplication by the rotation angle frees the norm and yields three parameter combinations that are fully independent: $\omega n_x, \omega n_y, \omega n_z$. The rotation matrices thus cover the full parameter space of the orthogonal matrices with determinant $+1$ and form a basis for the SO(3) group. By multiplying the rotation matrices with the negative unit matrix $-\mathbb{I}$, we obtain a result with determinant -1, which belongs to the coset of SO(3) matrices. The $-\mathbb{I}$ matrix represents the inversion of all points through the origin to their antipodes. When $\det\mathbb{A} = 1$, the corresponding rotation is called a *proper rotation* because it represents rotations as described earlier. When, on the other hand, $\det\mathbb{A} = -1$, the corresponding operation is a combination of a rotation and an inversion, and is called an *improper rotation*. In the following, attention will be focused on proper rotations only, with $\det\mathbb{R}(\omega\boldsymbol{n}) = 1$, described by the SO(3) group. This symmetry group is of paramount importance in theoretical physics because most classical and quantum systems have spherical symmetry.

The subset of rotations $\{\mathbb{R}(\omega\boldsymbol{n})\}$ around a *fixed* rotation axis forms a subgroup of the three-dimensional rotation group SO(3). This subgroup is Abelian because two successive rotations commute when they share the same rotation axis; that is, the product can be written as a single rotation about the axis through the angle $\omega_3 = \omega_1+\omega_2 = \omega_2+\omega_1$. This subgroup is the familiar rotation group in two dimensions of the previous chapter, denoted SO(2). The relationship between the Lie groups can be represented by a descending *chain of subgroups*:

$$\mathrm{O}(3) \supset \mathrm{SO}(3) \supset \mathrm{SO}(2). \tag{5.46}$$

5.3.4 The Euler parameterization

The SO(3) group contains an infinite number of rotation matrices $\mathbb{R}(\omega\boldsymbol{n})$ that are continuously dependent on *three parameters*. Every rotation needs to be specified, in other words, by three continuously varying parameters. The rotation group is therefore called an infinite continuous group of *dimension 3*, or a *triparametric Lie group*.

The idea that three-dimensional rotations necessitated the adoption of three parameters originated in the work of renowned mathematician Leonhard Euler. In 1775, Euler published a paper dealing with the mathematics of two-, three-, four- and five-dimensional rotations, in which he set forth his idea that three parameters were required to specify a three-dimensional rotation $\mathbb{R}(\omega \boldsymbol{n})$.[5] His statement became known as *Euler's rotation theorem* and was soon generalized to rotations in n dimensions:

Theorem 5.1 (Euler's generalized rotation theorem): Every n-dimensional rotation can be characterized by a minimum of $n(n-1)/2$ parameters. ■

Two-dimensional rotations are thus specified by $2(2-1)/2 = 1$ parameter; three-dimensional rotations by $3(3-1)/2 = 3$ parameters; and four-dimensional rotations by $4(4-1)/2 = 6$ parameters (as we shall see in Chapter 9).

As we saw in the preceding section, Euler's theorem derives directly from the skew-symmetry of the $\mathbb{S}$ matrix in the exponential form of the orthogonal matrices.

The Euler parameterization of SO(3) rotations is based on three consecutive rotations around two Cartesian directions, and hence on three angles. In the current treatment we have given preference to the expression in Eq. (5.45), which puts rotations around all three Cartesian directions on an equal footing.

5.4 THE $\mathfrak{so}(3)$ LIE ALGEBRA

5.4.1 The $\mathfrak{so}(3)$ generators

The Lie algebra, associated with the SO(3) group, can now be easily constructed. In SO(3), the role of the unique $\mathbb{X}$ generator in SO(2) is taken over by the $\mathbb{Z}$ generator, which has the form of a scalar product between the directional cosines of the rotation axis and three elementary generator matrices. Let us define these constituent generators as $\mathbb{X}_1$, $\mathbb{X}_2$, and $\mathbb{X}_3$, corresponding to n_x, n_y, and n_z, respectively:

$$\begin{aligned}\mathbb{Z} &= n_x\mathbb{X}_1 + n_y\mathbb{X}_2 + n_z\mathbb{X}_3 \\ &= n_x\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} + n_y\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + n_z\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned} \tag{5.47}$$

These generator components correspond to rotations around the respective Cartesian directions. As an example for $n_x = 1$, the rotation matrix becomes

$$\mathbb{R}(\omega\, \boldsymbol{e}_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix}. \tag{5.48}$$

The generator for this rotation is then obtained exactly as in the case of the SO(2) group, by differentiating this matrix with respect to ω taken at $\omega = 0$, and similarly for rotations around the y- and z-directions:

---

[5] Euler presented this result to the St. Petersburg Academy on October 9, 1775. It was published in: L. Euler. *Novi Commentarii Academiae Scientiarum in Petropolitanae* 20 (1776), pp. 189–207.

$$\begin{aligned}\mathbb{X}_1 = \frac{d}{d\omega}\mathbb{R}\left(\omega\, \boldsymbol{e}_x\right)\bigg|_{\omega=0} &= \frac{d}{d\omega}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix}\Bigg|_{\omega=0} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\omega & -\cos\omega \\ 0 & \cos\omega & -\sin\omega \end{bmatrix}\Bigg|_{\omega=0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix};\end{aligned} \tag{5.49}$$

$$\begin{aligned}\mathbb{X}_2 = \frac{d}{d\omega}\mathbb{R}\left(\omega\, \boldsymbol{e}_y\right)\bigg|_{\omega=0} &= \frac{d}{d\omega}\begin{bmatrix} \cos\omega & 0 & \sin\omega \\ 0 & 1 & 0 \\ -\sin\omega & 0 & \cos\omega \end{bmatrix}\Bigg|_{\omega=0} \\ &= \begin{bmatrix} -\sin\omega & 0 & \cos\omega \\ 0 & 0 & 0 \\ -\cos\omega & 0 & \sin\omega \end{bmatrix}\Bigg|_{\omega=0} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix};\end{aligned} \tag{5.50}$$

$$\begin{aligned}\mathbb{X}_3 = \frac{d}{d\omega}\mathbb{R}\left(\omega\, \boldsymbol{e}_z\right)\bigg|_{\omega=0} &= \frac{d}{d\omega}\begin{bmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{bmatrix}\Bigg|_{\omega=0} \\ &= \begin{bmatrix} -\sin\omega & -\cos\omega & 0 \\ \cos\omega & -\sin\omega & 0 \\ 0 & 0 & 0 \end{bmatrix}\Bigg|_{\omega=0} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned} \tag{5.51}$$

Let us list, by way of summary, the three generators of the SO(3) group in their matrix form:

$$\mathbb{X}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix};\ \mathbb{X}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix};\ \mathbb{X}_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.52}$$

These three matrices form the operators of our Lie algebra. To work out the structure of this algebra we have to consider their products. For an operator algebra the product combination rule is actually the commutator. This product rule is also known as the *Lie bracket*. Its appearance in the algebra of generators is explained in detail in Appendix F. We can verify immediately that, in the case of $\mathfrak{so}(3)$, the operators do *not* commute—that is,

$$\left[\mathbb{X}_i, \mathbb{X}_j\right] = \mathbb{X}_i\mathbb{X}_j - \mathbb{X}_j\mathbb{X}_i \neq 0 \qquad \forall i,j = 1,2,3;\ i \neq j. \tag{5.53}$$

The set of generators $\{\mathbb{X}_1, \mathbb{X}_2, \mathbb{X}_3\}$ is, however, *closed* under commutation. Let us examine this concept by deriving the commutator $[\mathbb{X}_1, \mathbb{X}_2]$ of $\mathbb{X}_1$ and $\mathbb{X}_2$. Since $[\mathbb{X}_1, \mathbb{X}_2] = \mathbb{X}_1\mathbb{X}_2 - \mathbb{X}_2\mathbb{X}_1$, we start by calculating the matrix products $\mathbb{X}_1\mathbb{X}_2$ and $\mathbb{X}_2\mathbb{X}_1$, respectively:

$$\mathbb{X}_1\mathbb{X}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \tag{5.54}$$

$$\mathbb{X}_2\mathbb{X}_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.55}$$

The commutator $[\mathbb{X}_1, \mathbb{X}_2]$ is then easily found to equal the third generator $\mathbb{X}_3$:

$$[\mathbb{X}_1,\mathbb{X}_2] = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbb{X}_3. \tag{5.56}$$

Similar results are obtained for every other possible commutator $\left[\mathbb{X}_i, \mathbb{X}_j\right]$ $(i, j = 1, 2, 3;$ $i \neq j)$:

$$\begin{aligned} [\mathbb{X}_1,\mathbb{X}_2] &= \mathbb{X}_3; \qquad & [\mathbb{X}_2,\mathbb{X}_1] &= -\mathbb{X}_3; \\ [\mathbb{X}_2,\mathbb{X}_3] &= \mathbb{X}_1; \qquad & [\mathbb{X}_3,\mathbb{X}_2] &= -\mathbb{X}_1; \\ [\mathbb{X}_3,\mathbb{X}_1] &= \mathbb{X}_2; \qquad & [\mathbb{X}_1,\mathbb{X}_3] &= -\mathbb{X}_2. \end{aligned} \tag{5.57}$$

This proves the closure under commutation. The *commutation relations* of the generators of the SO(3) group can be summarized as follows:

$$\left[\mathbb{X}_i, \mathbb{X}_j\right] = \sum_{k=1}^{3} \varepsilon_{ijk} \mathbb{X}_k \equiv \varepsilon_{ijk} \mathbb{X}_k, \tag{5.58}$$

where we have used Einstein's summation convention in the last term (i.e., repeated indices imply summation over the range of possible values of these indices), and where the permutation symbol ε_{ijk} denotes the elements of the *Levi-Civita antisymmetric three-dimensional tensor* (after the Italian mathematician and physicist Tullio Levi-Civita [1873–1941]):

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) = (1,2,3),\ (2,3,1) \text{ or } (3,1,2) \\ -1 & \text{if } (i,j,k) = (1,3,2),\ (2,1,3) \text{ or } (3,2,1). \\ 0 & \text{if } i = j,\ j = k \text{ or } k = i \end{cases} \tag{5.59}$$

The set of generators $\{\mathbb{X}_1, \mathbb{X}_2, \mathbb{X}_3\}$ and their commutation relations are said to define the *Lie algebra* $\mathfrak{so}(3)$ of the *Lie group* SO(3).

Definition 5.3 (The $\mathfrak{so}(3)$ Lie algebra): The Lie algebra $\mathfrak{so}(3)$ is a linear vector space, spanned by the generators $\mathbb{X}_i$ $(i = 1, 2, 3)$, that is combined with a multiplication operator (i.e., product rule) $[\cdot,\cdot] : \mathfrak{so}(3) \times \mathfrak{so}(3) \rightarrow \mathfrak{so}(3)$, called the *Lie bracket*, for which $\left[\mathbb{X}_i, \mathbb{X}_j\right] = \varepsilon_{ijk}\mathbb{X}_k$, with ε_{ijk} the so-called *structure constant*. ■

Because the generators $\mathbb{X}_i$ form a *basis* of the Lie algebra, every $\mathbb{X} \in \mathfrak{so}(3)$ can be written as a linear combination of the $\mathbb{X}_i$'s:

$$\mathbb{X} = \sum_{i=1}^{3} c_i \mathbb{X}_i \equiv c_i \mathbb{X}_i; \qquad c_i \in \mathbb{R}. \tag{5.60}$$

An example is the $\mathbb{Z}$ matrix defined in Eq. (5.47), which describes an infinitesimal rotation along the direction of the unit vector $\boldsymbol{n}$:

$$\mathbb{Z} = n_x \mathbb{X}_1 + n_y \mathbb{X}_2 + n_z \mathbb{X}_3. \tag{5.61}$$

5.4.2 Operator form of the SO(3) generators

Instead of formulating the three generators $\mathbb{X}_i$ $(i = 1, 2, 3)$ of the SO(3) group in terms of *matrices*, as in Eq. (5.52), we want to find their corresponding *operator* formulation. This reformulation in terms of differential operators enables us to form the bridge between the group theory of three-dimensional rotations (as described earlier) and the quantum mechanics of angular momentum (see §5.5). A deeper understanding of the *physical* role of the $\mathfrak{so}(3)$ generators will thus be achieved.

The derivation of this operator is obtained by a straightforward generalization of the operator for the rotation around the z-axis, as described in Chapter 4. Recall that the generator matrix and the corresponding operator form for a rotation in the xy-plane was given by:

$$\mathbb{X} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}; \tag{5.62}$$

$$\hat{X} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}. \tag{5.63}$$

In the SO(3) group, the generator matrix describing the infinitesimal rotation about the z-axis is denoted as $\mathbb{X}_3$. Because this rotation does not affect the z-coordinate, the operator form is precisely the same as obtained for the two-dimensional SO(2) group. The derivation of the operator form of the other two generators, $\mathbb{X}_1$ and $\mathbb{X}_2$, proceeds in a completely analogous way. By way of comparison, a list is provided here of all the generators $\mathbb{X}_i$ $(i = 1,2,3)$ in their respective matrix and operator forms:

$$\mathbb{X}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}; \; \mathbb{X}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}; \; \mathbb{X}_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \tag{5.64}$$

$$\hat{X}_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}; \quad \hat{X}_2 = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}; \quad \hat{X}_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}. \tag{5.65}$$

Note that these operators are related by a cyclic permutation of the Cartesian indices, as $x \to y \to z$.

5.5 ROTATIONS IN QUANTUM MECHANICS

5.5.1 Angular momentum as the generator of rotations

Having established the operator formulation of the $\mathfrak{so}(3)$ generators, we are now in a position to assign them a physical interpretation by making the comparison with the components $\hat{L}_i$ $(i = 1,2,3)$ of the *orbital angular momentum* $\hat{\mathbf{L}}$, as derived in Chapter 4:

$$\hat{L}_1 = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right); \hat{L}_2 = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right); \hat{L}_3 = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \tag{5.66}$$

On comparison, the components $\hat{L}_i$ of the orbital angular momentum are seen to correspond to the generators $\hat{X}_i$ of the $\mathfrak{so}(3)$ algebra via the equation

$$\hat{L}_i = i\hbar\hat{X}_i; \qquad i = 1,2,3. \tag{5.67}$$

In the words of José Alvarado, Eq. (5.67) “follows the convention QM = $i\hbar$ GT. The factor $i\hbar$ provides the link between group theoretical generators, which are produced by Lie groups and Lie algebras, and quantum mechanical operators, which correspond to physically meaningful quantities.”[6] We can, in other words, *identify* the components of the orbital angular momentum with the generators of the SO(3)

---

[6] J. Alvarado. *Group Theoretical Aspects of Quantum Mechanics.* 2007, p. 45. Script based on the Group Theory seminar by Prof. G. Rudolph during the winter semester 2004–2005 at the University of Leipzig, see: www.alv.ac/physik/Group-Theory.pdf.

group (at least, up to a factor $i\hbar$). This explains why we generally say that the $\hat{L}_i$ *components generate infinitesimal rotations around the mutually orthogonal* x, y, *and* z*-axes in three-dimensional space.*

5.5.2 The rotation operator

We can elaborate on this line of thought by introducing a rotation operator, $\hat{U}(\omega\boldsymbol{n})$, associated with the matrix transformation $\mathbb{R}(\omega\boldsymbol{n}) = e^{\omega\boldsymbol{n}\cdot\mathbb{Z}}$. This operator can be expressed in terms of the angular momentum components $\hat{L}_i$. Rewriting Eq. (5.67) as

$$\hat{X}_i = \frac{1}{i\hbar}\hat{L}_i = -\frac{i}{\hbar}\hat{L}_i; \qquad i = 1,2,3, \tag{5.68}$$

and substituting this expression in Eq. (5.41) yields

$$\begin{aligned}\hat{U}(\omega\boldsymbol{n}) &= e^{\omega n_x \hat{X}_1 + \omega n_y \hat{X}_2 + \omega n_z \hat{X}_3} \\ &= e^{-(i/\hbar)\omega n_x \hat{L}_1 - (i/\hbar)\omega n_y \hat{L}_2 - (i/\hbar)\omega n_z \hat{L}_3} \\ &= e^{-(i/\hbar)\omega\left[n_x \hat{L}_1 + n_y \hat{L}_2 + n_z \hat{L}_3\right]} \\ &= e^{-(i/\hbar)\omega\boldsymbol{n}\cdot\hat{\mathbf{L}}}.\end{aligned} \tag{5.69}$$

This equation enables us to define the quantum mechanical rotation operator.

Definition 5.4 (Rotation operator): To every three-dimensional rotation $\hat{R}$, represented by the rotation matrix $\mathbb{R}(\omega\boldsymbol{n})$, corresponds a unique *rotation operator*

$$\hat{U}(\omega\boldsymbol{n}) \equiv e^{-(i/\hbar)\omega\boldsymbol{n}\cdot\hat{\mathbf{L}}}, \tag{5.70}$$

with $\omega\boldsymbol{n} = \begin{bmatrix} \omega n_x & \omega n_y & \omega n_z \end{bmatrix}^{\mathrm{T}}$ and $\hat{\mathbf{L}} = \begin{bmatrix} \hat{L}_1 & \hat{L}_2 & \hat{L}_3 \end{bmatrix}^{\mathrm{T}}$. ∎

5.6 ANGULAR MOMENTUM

5.6.1 The angular momentum algebra

With the help of Eq. (5.67), the commutation relations of the $\mathfrak{so}(3)$ generators $\hat{X}_i$ in Eq. (5.58) translate into

$$\left[\hat{L}_i, \hat{L}_j\right] = i\hbar \sum_{k=1}^{3} \varepsilon_{ijk}\hat{L}_k \equiv i\hbar\varepsilon_{ijk}\hat{L}_k, \tag{5.71}$$

which are the familiar commutation relations of the orbital angular momentum. Eq. (5.71) defines the *angular momentum algebra*. As will be illustrated in the next few paragraphs, most of the properties of angular momentum (such as the allowed spectrum of angular momentum eigenstates and eigenvalues) simply follow from this algebra—and this in a most elegant way.

5.6.2 Casimir invariants

To determine the consequences of the angular momentum algebra, we start by introducing the $\hat{L}^2$ operator for the square of the total angular momentum:

$$\hat{L}^2 = \hat{\mathbf{L}}\cdot\hat{\mathbf{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \tag{5.72}$$

which has the interesting property that it commutes with all three angular momentum operators:

$$\left[\hat{L}^2, \hat{L}_i\right] = 0, \qquad i = 1,2,3. \tag{5.73}$$

To prove this, we use the commutation relations of Eq. (5.71) and calculate

$$\begin{aligned}
\left[\hat{L}^2, \hat{L}_x\right] &= \left[\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x\right] \\
&= \left[\hat{L}_x^2, \hat{L}_x\right] + \left[\hat{L}_y^2, \hat{L}_x\right] + \left[\hat{L}_z^2, \hat{L}_x\right] \\
&= \left[\hat{L}_y^2, \hat{L}_x\right] + \left[\hat{L}_z^2, \hat{L}_x\right] \\
&= \left[\hat{L}_y, \hat{L}_x\right]\hat{L}_y + \hat{L}_y\left[\hat{L}_y, \hat{L}_x\right] + \left[\hat{L}_z, \hat{L}_x\right]\hat{L}_z + \hat{L}_z\left[\hat{L}_z, \hat{L}_x\right] \\
&= -i\hbar\hat{L}_z\hat{L}_y - i\hbar\hat{L}_y\hat{L}_z + i\hbar\hat{L}_y\hat{L}_z + i\hbar\hat{L}_z\hat{L}_y \\
&= 0.
\end{aligned} \tag{5.74}$$

Since a cyclic permutation of x, y, and z leaves $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ invariant, we can carry out two such permutations on Eq. (5.74) to obtain

$$\left[\hat{L}^2, \hat{L}_y\right] = 0 \quad \text{and} \quad \left[\hat{L}^2, \hat{L}_z\right] = 0. \tag{5.75}$$

When an operator such as $\hat{L}^2$ commutes (by virtue of its nature) with every generator $\hat{L}_i$ of the $\mathfrak{so}(3)$ algebra, it is called a *Casimir operator* or a *Casimir invariant* of that algebra.[7]

5.6.3 The eigenvalue problem

In view of its commuting properties, $\hat{L}^2$ shares a complete set of simultaneous eigenfunctions with any one of the $\hat{L}_i$ operators. However, because the $\hat{L}_i$ operators do *not* commute among themselves, we have to agree on a *quantization scheme*. In other words, we have to select *one* (and only one) angular momentum operator component that henceforth characterizes the eigenstates (sometimes also called *substates*) inside the shell.[8] The standard choice corresponds to the $\hat{L}_3 = \hat{L}_z$ operator (although we might just as well have chosen the operator $\hat{L}_x$ or $\hat{L}_y$). We say that $\hat{L}^2$ and $\hat{L}_z$ can be *diagonalized* simultaneously.

The angular momentum states are thus simultaneous eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$. Representing these states as $|\alpha,\beta\rangle$, we have

$$\hat{L}^2\,|\alpha,\beta\rangle = \alpha\,|\alpha,\beta\rangle; \tag{5.76}$$

$$\hat{L}_z\,|\alpha,\beta\rangle = \beta\,|\alpha,\beta\rangle, \tag{5.77}$$

where α and β are the eigenvalues of $\hat{L}^2$ and $\hat{L}_z$, respectively. For the time being, both the eigenkets and their corresponding eigenvalues are unknown. This is called

---

[7] The Casimir invariant is named after Dutch physicist Hendrik Casimir (1909–2000).

[8] The fact that the components of the angular momentum do not share a set of common eigenfunctions is an example of Heisenberg's *uncertainty principle*, which says that $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ are *incompatible* observables. See also Appendix D, §D.3.

the *eigenvalue problem* of quantum mechanics. One particularly powerful way of solving the eigenvalue problem for α and β involves the clever exploitation of the commutation relations in Eq. (5.71) by the use of so-called *ladder operators* (the reason for this terminology will become clear in the next section).

5.6.4 Dirac's ladder operator method

The introduction of ladder operators is usually credited to British theoretical physicist Paul A. M. Dirac (1902–1984), who first used them to solve the angular momentum wave equations. The *ladder operator method* is an algebraic technique of surprising simplicity that exploits the commutation relations of the angular momentum operators fully. Our knowledge of the $\mathfrak{so}(3)$ Lie algebra will, in other words, prove to be more than sufficient to determine the possible eigenkets $|\alpha, \beta\rangle$ and eigenvalues α and β of $\hat{L}^2$ and $\hat{L}_z$.

General formulation

The eigenvalue problem of quantum mechanics consists of finding the eigenkets $|a_i\rangle$ and eigenvalues a_i of a Hermitian operator $\hat{A}$ with the eigenvalue equation

$$\hat{A}\,|a_i\rangle = a_i\,|a_i\rangle. \tag{5.78}$$

The ladder operator method of Dirac makes use of mathematical *raising* and *lowering operators*, which are collectively known as *ladder operators.*[9] A ladder operator $\hat{B}$ is an operator that transforms the eigenkets of $\hat{A}$ into new eigenkets of $\hat{A}$, with the eigenvalue a_i being either raised or lowered by a certain constant value c (Figure 5.2). That is,

$$\hat{A}\hat{B}|a_i\rangle = (a_i + c)\,\hat{B}\,|a_i\rangle. \tag{5.79}$$

Thus, if $|a_i\rangle$ is an eigenket of $\hat{A}$ with eigenvalue a_i, then $\hat{B}|a_i\rangle$ represents another eigenket of $\hat{A}$ with eigenvalue $a_i + c$. The operator $\hat{B}$ is said to be a *raising operator* for $\hat{A}$ when the constant c is real and positive, and it is called a *lowering operator* for $\hat{A}$ when c is real but negative.

An inquiry into the nature of $\hat{B}$ has led to the formulation of an interesting theorem that lists the fundamental *operator relation* between $\hat{A}$ and $\hat{B}$ that should be satisfied for $\hat{B}$ to behave as a ladder operator.

Theorem 5.2 (Ladder operators): An operator $\hat{B}$ functions as a ladder operator for the eigenvalues of a Hermitian operator $\hat{A}$ whenever its commutator with $\hat{A}$ satisfies the

---

[9] Some authors prefer the term *step operators* or *shift operators*. Raising operators are also called *creation operators* whereas lowering operators are termed *annihilation* (or *destruction*) *operators* for reasons that will become obvious in Chapter 7.

FIGURE 5.2 Let Alice and Humpty Dumpty represent two states $|A\rangle$ and $|HD\rangle$ of the system, with different eigenvalues (heights). With the help of *ladder operators*, Alice can either climb up to Humpty Dumpty, or Humpty Dumpty can make his great fall. That is, if a *raising operator* $\hat{B}_+$ acts on the state $|A\rangle$, a new state $|HD\rangle$ is created with a raised eigenvalue: $\hat{B}_+ |A\rangle \to |HD\rangle$. The opposite transformation is effected under the action of a *lowering operator* $\hat{B}_-$, which acts on $|HD\rangle$ to yield $|A\rangle$ with a lowered eigenvalue: $\hat{B}_- |HD\rangle \to |A\rangle$.

operator relation

$$\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A} = c\hat{B}, \tag{5.80}$$

for some nonzero real scalar c. ■

Proof. Given the eigenvalue equation for $\hat{A}$ in Eq. (5.78), it is clear that the operator $\hat{B}$ acts on $|a_i\rangle$ in such a way as to shift the eigenvalue by the amount c:

$$\begin{aligned}
\hat{A}\hat{B}|a_i\rangle &= \left(\hat{B}\hat{A} + \left[\hat{A}, \hat{B}\right]\right)|a_i\rangle \\
&= \left(\hat{B}\hat{A} + c\hat{B}\right)|a_i\rangle \\
&= \hat{B}\hat{A}|a_i\rangle + c\hat{B}|a_i\rangle \\
&= \hat{B}a_i|a_i\rangle + c\hat{B}|a_i\rangle \\
\hat{A}\hat{B}|a_i\rangle &= (a_i + c)\,\hat{B}|a_i\rangle,
\end{aligned} \tag{5.81}$$

where use was made of Eq. (5.80). ■

Just as $|a_i\rangle$ is called an *eigenfunction* of $\hat{A}$ with *eigenvalue* a_i in view of Eq. (5.78), $\hat{B}$ is said to be an *eigenoperator* of $\hat{A}$ with *eigenvalue* c in view of Eq. (5.80).

The angular momentum ladder operators

The *raising operator* $\hat{L}_+$ and *lowering operator* $\hat{L}_-$ for the eigenvalues of the angular momentum operator $\hat{L}_z$ turn out to have the following forms:

$$\hat{L}_+ \equiv \hat{L}_x + i\hat{L}_y; \tag{5.82}$$

$$\hat{L}_- \equiv \hat{L}_x - i\hat{L}_y. \tag{5.83}$$

Notice that $\hat{L}_+$ and $\hat{L}_-$ are the *adjoint* of each other; that is, $\hat{L}_- = \left(\hat{L}_+\right)^\dagger$. This implies that neither of these operators is *Hermitian* and they do not, in any way, correspond to physical observables, in contrast to the $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, and $\hat{L}^2$ operators.[10]

Both $\hat{L}_+$ and $\hat{L}_-$ should satisfy Eq. (5.80) of Theorem 5.2 to function as ladder operators. For this, we must evaluate the commutators of the raising and lowering operators with $\hat{L}_z$. Using the commutation relations in Eq. (5.71) for the angular momentum operators, we obtain

$$\begin{aligned}\left[\hat{L}_z, \hat{L}_+\right] &= \left[\hat{L}_z, \hat{L}_x + i\hat{L}_y\right] \\ &= \left[\hat{L}_z, \hat{L}_x\right] + i\left[\hat{L}_z, \hat{L}_y\right] \\ &= i\hbar\hat{L}_y + \hbar\hat{L}_x \\ &= \hbar\left(\hat{L}_x + i\hat{L}_y\right) \\ \left[\hat{L}_z, \hat{L}_+\right] &= \hbar\hat{L}_+.\end{aligned} \tag{5.84}$$

Similarly, we find

$$\left[\hat{L}_z, \hat{L}_-\right] = -\hbar\hat{L}_-. \tag{5.85}$$

Both $\hat{L}_+$ and $\hat{L}_-$ are seen to satisfy the operator relation of Theorem 5.2, which implies that both are, indeed, ladder operators.

Operating with $\hat{L}_+$ and $\hat{L}_-$ on the eigenkets of $\hat{L}_z$

Let us act with the raising operator on the eigenket $|\alpha, \beta\rangle$ and determine the outcome of $\hat{L}_z\hat{L}_+ |\alpha, \beta\rangle$. We start by rewriting Eqs. (5.84) and (5.85) as

$$\hat{L}_z\hat{L}_+ = \hat{L}_+\hat{L}_z + \hbar\hat{L}_+; \tag{5.86}$$

$$\hat{L}_z\hat{L}_- = \hat{L}_-\hat{L}_z - \hbar\hat{L}_-. \tag{5.87}$$

Armed with these results, and the fact that $\hat{L}_+$ is linear, we have

$$\begin{aligned}\hat{L}_z\hat{L}_+ |\alpha, \beta\rangle &= \left(\hat{L}_+\hat{L}_z + \hbar\hat{L}_+\right)|\alpha, \beta\rangle \\ &= \hat{L}_+\hat{L}_z |\alpha, \beta\rangle + \hbar\hat{L}_+ |\alpha, \beta\rangle \\ &= \hat{L}_+\beta |\alpha, \beta\rangle + \hbar\hat{L}_+ |\alpha, \beta\rangle \\ &= (\beta + \hbar)\,\hat{L}_+ |\alpha, \beta\rangle.\end{aligned} \tag{5.88}$$

---

[10] The $\hat{L}_x$ and $\hat{L}_y$ operators can be written in terms of the ladder operators:

$$\hat{L}_x = \frac{1}{2}\left(\hat{L}_+ + \hat{L}_-\right) \quad \text{and} \quad \hat{L}_y = \frac{1}{2i}\left(\hat{L}_+ - \hat{L}_-\right).$$

According to this eigenvalue equation, the function $\hat{L}_+ |\alpha,\beta\rangle$ is an eigenket of $\hat{L}_z$ with an eigenvalue $\beta+\hbar$. That is to say, when we operate with the raising operator $\hat{L}_+$ on the eigenket $|\alpha,\beta\rangle$, the eigenket is transformed into another eigenket of $\hat{L}_z$ with an eigenvalue that has been raised by 1 unit of $\hbar$ in comparison with the eigenvalue of $|\alpha,\beta\rangle$. Applying the raising operator $\hat{L}_+$ once more to Eq. (5.88) yields, with the help of Eq. (5.86),

$$\hat{L}_z\hat{L}_+^2 |\alpha,\beta\rangle = (\beta+2\hbar)\,\hat{L}_+^2 |\alpha,\beta\rangle. \tag{5.89}$$

Generalizing for the k-fold application of $\hat{L}_+$ on $|\alpha,\beta\rangle$,

$$\hat{L}_z\hat{L}_+^k |\alpha,\beta\rangle = (\beta+k\hbar)\,\hat{L}_+^k |\alpha,\beta\rangle, \qquad k=0,1,2,\ldots. \tag{5.90}$$

We can also operate on the eigenvalue equation (5.76) with the lowering operator $\hat{L}_-$. Using Eq. (5.87), and proceeding in complete analogy, leads to

$$\hat{L}_z\hat{L}_- |\alpha,\beta\rangle = (\beta-\hbar)\,\hat{L}_- |\alpha,\beta\rangle; \tag{5.91}$$

$$\hat{L}_z\hat{L}_-^k |\alpha,\beta\rangle = (\beta-k\hbar)\,\hat{L}_-^k |\alpha,\beta\rangle, \qquad k=0,1,2,\ldots. \tag{5.92}$$

It follows from Eqs. (5.90) and (5.92) that the functions $\hat{L}_\pm^k |\alpha,\beta\rangle$ are eigenkets of $\hat{L}_z$ with the eigenvalues $\beta\pm k\hbar$:

$$\hat{L}_z\hat{L}_\pm^k |\alpha,\beta\rangle = (\beta\pm k\hbar)\,\hat{L}_\pm^k |\alpha,\beta\rangle, \qquad k=0,1,2,\ldots. \tag{5.93}$$

We can, in other words, generate a *ladder* of eigenvalues by the repeated application of the raising and lowering operators on the eigenket $|\alpha,\beta\rangle$ with the eigenvalue β:

$$\ldots \xleftrightarrow{\pm\hbar} \beta-2\hbar \xleftrightarrow{\pm\hbar} \beta-\hbar \xleftrightarrow{\pm\hbar} \beta \xleftrightarrow{\pm\hbar} \beta+\hbar \xleftrightarrow{\pm\hbar} \beta+2\hbar \xleftrightarrow{\pm\hbar} \ldots, \tag{5.94}$$

where the eigenvalues differ by integral multiples of $\hbar$.

Operating with $\hat{L}_+$ and $\hat{L}_-$ on the eigenkets of $\hat{L}^2$

We now show that the functions $\hat{L}_\pm^k |\alpha,\beta\rangle$ are also eigenkets of $\hat{L}^2$, all with the *same* eigenvalue α:

$$\hat{L}^2\hat{L}_\pm^k |\alpha,\beta\rangle = \alpha\hat{L}_\pm^k |\alpha,\beta\rangle, \qquad k=0,1,2,\ldots. \tag{5.95}$$

To prove Eq. (5.95), we first show that $\hat{L}^2$ commutes with the ladder operators $\hat{L}_+$ and $\hat{L}_-$. Using Eqs. (5.74) and (5.75):

$$\left[\hat{L}^2,\hat{L}_\pm\right] = \left[\hat{L}^2,\hat{L}_x \pm i\hat{L}_y\right] = \left[\hat{L}^2,\hat{L}_x\right] \pm i\left[\hat{L}^2,\hat{L}_y\right] = 0. \tag{5.96}$$

We also have

$$\left[\hat{L}^2,\hat{L}_\pm^2\right] = \left[\hat{L}^2,\hat{L}_\pm\right]\hat{L}_\pm + \hat{L}_\pm\left[\hat{L}^2,\hat{L}_\pm\right] = 0, \tag{5.97}$$

and it follows by induction that

$$\left[\hat{L}^2,\hat{L}_\pm^k\right] = 0 \quad \text{or} \quad \hat{L}^2\hat{L}_\pm^k = \hat{L}_\pm^k\hat{L}^2, \qquad k=0,1,2,\ldots. \tag{5.98}$$

If we operate on Eq. (5.76) with $\hat{L}_\pm^k$ and use Eq. (5.98), we get

$$\hat{L}^2\hat{L}_\pm^k |\alpha,\beta\rangle = \hat{L}_\pm^k\hat{L}^2 |\alpha,\beta\rangle = \hat{L}_\pm^k\alpha |\alpha,\beta\rangle = \alpha\hat{L}_\pm^k |\alpha,\beta\rangle, \tag{5.99}$$

which is what we wanted to prove.

Determining the eigenvalues of $\hat{L}_z$ and $\hat{L}^2$

Next, we show that the set of eigenvalues of $\hat{L}_z$, generated with the ladder operators $\hat{L}_+$ and $\hat{L}_-$, must be *bounded* from above and below. For the particular eigenket $|\alpha,\beta\rangle$ with $\hat{L}_z$ eigenvalue β, we have

$$\hat{L}_z\,|\alpha,\beta\rangle = \beta\,|\alpha,\beta\rangle, \tag{5.100}$$

and for the set of eigenkets and eigenvalues generated by the ladder operators, we have

$$\hat{L}_z\,|\alpha,\beta_k\rangle = \beta_k\,|\alpha,\beta_k\rangle, \tag{5.101}$$

where

$$|\alpha,\beta_k\rangle = \hat{L}_\pm^k\,|\alpha,\beta\rangle; \tag{5.102}$$

$$\beta_k = \beta \pm k\hbar. \tag{5.103}$$

Operating on Eq. (5.101) with $\hat{L}_z$, we have

$$\begin{aligned}\hat{L}_z^2\,|\alpha,\beta_k\rangle &= \beta_k\hat{L}_z\,|\alpha,\beta_k\rangle;\\ \hat{L}_z^2\,|\alpha,\beta_k\rangle &= \beta_k^2\,|\alpha,\beta_k\rangle.\end{aligned} \tag{5.104}$$

Now subtract Eq. (5.104) from Eq. (5.95), and use Eqs. (5.72) and (5.102):

$$\begin{aligned}\hat{L}^2\,|\alpha,\beta_k\rangle - \hat{L}_z^2\,|\alpha,\beta_k\rangle &= \alpha\,|\alpha,\beta_k\rangle - \beta_k^2\,|\alpha,\beta_k\rangle\\ \left(\hat{L}_x^2+\hat{L}_y^2\right)|\alpha,\beta_k\rangle &= \left(\alpha-\beta_k^2\right)|\alpha,\beta_k\rangle.\end{aligned} \tag{5.105}$$

The operator $\hat{L}_x^2+\hat{L}_y^2$ corresponds to a nonnegative physical quantity and hence has nonnegative eigenvalues. Therefore, Eq. (5.105) implies that $\alpha-\beta_k^2 \geqslant 0$ and $\alpha^{1/2} \geqslant |\beta_k|$. Thus,

$$-\alpha^{1/2} \leqslant \beta_k \leqslant \alpha^{1/2}, \qquad k=0,\pm1,\pm2,\ldots. \tag{5.106}$$

Since α remains constant as k varies, Eq. (5.106) shows that the set of eigenvalues β_k is bounded above and below. Let $\beta_{\max}$ and $\beta_{\min}$ denote the maximum and minimum values of β_k, with $|\alpha,\beta_{\max}\rangle$ and $|\alpha,\beta_{\min}\rangle$ as the respective eigenkets:

$$\hat{L}_z\,|\alpha,\beta_{\max}\rangle = \beta_{\max}\,|\alpha,\beta_{\max}\rangle; \tag{5.107}$$

$$\hat{L}_z\,|\alpha,\beta_{\min}\rangle = \beta_{\min}\,|\alpha,\beta_{\min}\rangle. \tag{5.108}$$

Now operate on Eq. (5.107) with the raising operator and use Eq. (5.86):

$$\begin{aligned}\hat{L}_+\hat{L}_z\,|\alpha,\beta_{\max}\rangle &= \beta_{\max}\hat{L}_+\,|\alpha,\beta_{\max}\rangle\\ \hat{L}_z\hat{L}_+\,|\alpha,\beta_{\max}\rangle &= (\beta_{\max}+\hbar)\,\hat{L}_+\,|\alpha,\beta_{\max}\rangle.\end{aligned} \tag{5.109}$$

This last equation seems to contradict the statement that $\beta_{\max}$ is the largest eigenvalue of $\hat{L}_z$, because it says that $L_+\,|\alpha,\beta_{\max}\rangle$ is an eigenfunction of $\hat{L}_z$ with eigenvalue $\beta_{\max}+\hbar$. The only way out of this contradiction is to have $L_+\,|\alpha,\beta_{\max}\rangle$ vanish. (We always reject zero as an eigenfunction on physical grounds.) Thus,

$$\hat{L}_+\,|\alpha,\beta_{\max}\rangle = 0. \tag{5.110}$$

Let us now operate on Eq. (5.110) with the lowering operator. This requires us to elaborate the combination of raising and lowering operators, which again can be resolved by using only the Lie commutator algebra. We have

$$\begin{aligned}\hat{L}_-\hat{L}_+ &= \left(\hat{L}_x - i\hat{L}_y\right)\left(\hat{L}_x + i\hat{L}_y\right) \\ &= \hat{L}_x\left(\hat{L}_x + i\hat{L}_y\right) - i\hat{L}_y\left(\hat{L}_x + i\hat{L}_y\right) \\ &= \hat{L}_x^2 + i\hat{L}_x\hat{L}_y - i\hat{L}_y\hat{L}_x + \hat{L}_y^2 \\ &= \hat{L}^2 - \hat{L}_z^2 - i\left[\hat{L}_y, \hat{L}_x\right] \\ \hat{L}_-\hat{L}_+ &= \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z.\end{aligned} \tag{5.111}$$

Similarly, we find that

$$\hat{L}_+\hat{L}_- = \hat{L}^2 - \hat{L}_z^2 + \hbar\hat{L}_z. \tag{5.112}$$

Using Eqs. (5.110) and (5.111), and operating with $\hat{L}_-$ on $L_+ |\alpha, \beta_{\max}\rangle$ gives

$$\begin{aligned}\hat{L}_-\hat{L}_+ |\alpha, \beta_{\max}\rangle &= \left(\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z\right)|\alpha, \beta_{\max}\rangle \\ &= \left(\alpha - \beta_{\max}^2 - \hbar\beta_{\max}\right)|\alpha, \beta_{\max}\rangle \\ &= 0.\end{aligned} \tag{5.113}$$

This holds only when

$$\begin{aligned}\alpha - \beta_{\max}^2 - \hbar\beta_{\max} &= 0; \\ \alpha &= \beta_{\max}^2 + \hbar\beta_{\max}.\end{aligned} \tag{5.114}$$

A similar argument shows that

$$\hat{L}_- |\alpha, \beta_{\min}\rangle = 0, \tag{5.115}$$

and by applying the raising operator to this equation and using Eq. (5.112), we find

$$\alpha = \beta_{\min}^2 - \hbar\beta_{\min}. \tag{5.116}$$

Combining this last equation with Eq. (5.114), yields

$$\beta_{\max}^2 + \hbar\beta_{\max} = \beta_{\min}^2 - \hbar\beta_{\min}. \tag{5.117}$$

This is a quadratic equation relating the unknowns $\beta_{\max}$ and $\beta_{\min}$, and using the usual formula for quadratic equations (it still works in quantum mechanics), we find two roots:

$$\beta_{\max} = -\beta_{\min}; \quad \beta_{\max} = \beta_{\min} - \hbar. \tag{5.118}$$

The second root is rejected, because it says $\beta_{\max}$ is less than $\beta_{\min}$. So,

$$\beta_{\max} = -\beta_{\min}. \tag{5.119}$$

Moreover, Eq. (5.103) says $\beta_{\max}$ and $\beta_{\min}$ differ by an integral multiple of $\hbar$:

$$\beta_{\max} - \beta_{\min} = n\hbar, \qquad n = 0, 1, 2, \ldots. \tag{5.120}$$

Substituting Eq. (5.119) in Eq. (5.120) gives

$$\beta_{\max} = \frac{n}{2}\hbar, \qquad n = 0, 1, 2, \ldots, \tag{5.121}$$

and setting $j = n/2$ yields

$$\beta_{\max} = j\hbar, \qquad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots. \tag{5.122}$$

Similarly,

$$\beta_{\min} = -j\hbar, \qquad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots. \tag{5.123}$$

We thus obtain the following bounded set of possible $\hat{L}_z$ eigenvalues

$$\beta = -j\hbar,\ (-j+1)\,\hbar,\ (-j+2)\,\hbar,\ \ldots,\ (j-2)\,\hbar,\ (j-1)\,\hbar,\ j\hbar, \tag{5.124}$$

where j is either *integral* (for even n) or *half-integral* (for odd n). And from Eq. (5.114), we find as the $\hat{L}^2$ eigenvalues

$$\alpha = j(j+1)\,\hbar^2, \qquad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots. \tag{5.125}$$

After rewriting the eigenkets $|\alpha,\beta\rangle$ as $|j, m_j\rangle$, the following eigenvalue equations are obtained:

$$\hat{L}^2\,|j,m_j\rangle = j(j+1)\,\hbar^2\,|j,m_j\rangle, \qquad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots; \tag{5.126}$$

$$\hat{L}_z\,|j,m_j\rangle = m_j\hbar\,|j,m_j\rangle, \qquad m_j = -j, -j+1, \ldots, j-1, j. \tag{5.127}$$

The eigenket $|j, m_j\rangle$ is said to be an angular momentum eigenstate with angular momentum j and a z-component of the angular momentum m_j. It should be emphasized once more that the eigenvalues of $\hat{L}^2$ and $\hat{L}_z$ were found from the angular momentum commutation relations only. Thus, the quantization has been derived completely from the $\mathfrak{so}(3)$ Lie algebra and the requirement that physical observables be real. This is a remarkable result, to say the least; it shows the power of group theoretical approaches in quantum mechanics.

5.7 APPLICATION: PARTICLE ON A SPHERE

An atom has spherical symmetry, so since our aim is to understand the shell structure of atomic matter, we must leave the circle behind and consider problems with spherical symmetry. The simplest one is that of a single particle, which is confined to a spherical surface of constant radius. Simple as it may be, the *particle-on-a-sphere* model is very useful to describe the electronic structure of globular molecules. To derive the eigenvalues and eigenstates of the particle-on-a-sphere problem, we need to determine the Hamiltonian of a spherically symmetric system.

5.7.1 Spherical components of the Hamiltonian

The energy of a free particle with mass m moving in a uniform potential, which can be taken as the zero of energy, is purely kinetic. The corresponding Hamiltonian is given by

$$\begin{aligned}\hat{\mathscr{H}} &= \frac{1}{2m}\hat{p}^2 \\ &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right).\end{aligned} \tag{5.128}$$

In this equation, the momentum and hence the kinetic energy operator is decomposed into the standard Cartesian components, representing motions in three orthogonal directions. In a spherical problem, it is convenient to switch to radial and angular motions that are, respectively, along and orthogonal to the radius. We are thus looking for a spherical decomposition of the Laplacian:

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_{ang} + \hat{\mathscr{H}}_{rad}. \tag{5.129}$$

Instead of replacing the Cartesian coordinates with spherical ones to find both these components, we look for an alternative expression based entirely on position and momentum operators. The advantage of this formalism will become clear later on when we discuss the radial wave equation in Chapter 11.

Angular part of the Hamiltonian

From the previous chapter, we are already familiar with the angular momentum **L** describing the revolution of a particle around a central point:

$$|\mathbf{L}| = \left|\mathbf{r}\times\mathbf{p}\right| = |\mathbf{r}|\left|\mathbf{p}_{\perp}\right| = rp_{\perp}, \tag{5.130}$$

Here, $p_{\perp}$ denotes the momentum associated with the motion perpendicular to the radius. It follows that

$$p_{\perp} = \frac{|\mathbf{L}|}{r}; \quad p_{\perp}^2 = \frac{\mathbf{L}^2}{r^2} = \frac{L^2}{r^2}. \tag{5.131}$$

In classical physics, the kinetic energy corresponding to this motion is given by

$$E_{kin} = \frac{mv_{\perp}^2}{2} = \frac{p_{\perp}^2}{2m} = \frac{L^2}{2mr^2}. \tag{5.132}$$

In quantum physics, the angular momentum is replaced by its operator form. We then immediately obtain the angular part of the kinetic energy operator:

$$\hat{\mathscr{H}}_{ang} = \frac{1}{2m\hat{r}^2}\hat{L}^2. \tag{5.133}$$

Let us express this operator square in simple scalar products of r and p. In classical physics, this is straightforward:

$$\begin{aligned} L^2 &= r^2p^2\sin^2\alpha \\ &= r^2p^2 - r^2p^2\cos^2\alpha \\ &= r^2p^2 - (\mathbf{r}\cdot\mathbf{p})^2, \end{aligned} \tag{5.134}$$

where α is the angle between the vectors **r** and **p**. In view of the commutativity of the scalar product of vectors in classical physics (i.e., $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$), this equation is equivalent to

$$L^2 = r^2p^2 - (\mathbf{p}\cdot\mathbf{r})^2. \tag{5.135}$$

However, when replacing this expression with its operator form, an unexpected difficulty arises. The operator products $(\mathbf{r}\cdot\mathbf{p})$ and $(\mathbf{p}\cdot\mathbf{r})$ are not the same, because the position and momentum operators do not commute. In Appendix G, the explicit differential operator forms of both the left-hand and right-hand side of Eq. (5.135) are derived. The results are provided in Eqs. (G.7) and (G.10), respectively. Subtraction of both terms gives

$$\hat{L}^2 - \left(\hat{r}^2\hat{p}^2 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})^2\right) = \hbar^2\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right) = i\hbar\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right). \tag{5.136}$$

This enables us to reexpress $\hat{L}^2$ in terms of $\hat{r}$ and $\hat{p}$:

$$\hat{L}^2 = \hat{r}^2\hat{p}^2 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})^2 + i\hbar\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right). \tag{5.137}$$

Compared with the classical expression in Eq. (5.134), there is one additional term resulting from the noncommutation of position and momentum operators. The angular part of the Hamiltonian is thus given by

$$\hat{\mathscr{H}}_{ang} = \frac{1}{2m}\hat{p}^2 - \frac{1}{2m\hat{r}^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 + \frac{i\hbar}{2m\hat{r}^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right). \tag{5.138}$$

Radial part of the Hamiltonian

To obtain the radial part of the Hamiltonian, we must first define the *radial momentum operator*. Following the quantum mechanical recipe in Eq. (D.20), the radial momentum conjugate to the radial coordinate would simply be expressed as $\frac{\hbar}{i}\frac{\partial}{\partial r}$. This differential operator can be rewritten in Cartesian form using the chain rule and, consequently, can be simplified by introducing $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$:

$$\begin{aligned}\frac{\hbar}{i}\frac{\partial}{\partial r} &= \frac{\hbar}{i}\left(\frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} + \frac{\partial z}{\partial r}\frac{\partial}{\partial z}\right) \\ &= \frac{\hbar}{i}\frac{1}{r}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right) \\ &= \frac{1}{r}\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}.\end{aligned} \tag{5.139}$$

Here again we meet the difficulty that the expression depends on a product of operators that do not commute. We can overcome this difficulty by defining the true radial momentum operator as the average of this product and its reverse. Note that, in this way, the true momentum operator is real in the operator sense, as explained by Dirac[11]:

$$\hat{p}_r = \frac{1}{2}\left(\frac{1}{r}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} + \hat{\mathbf{p}}\cdot\hat{\mathbf{r}}\,\frac{1}{r}\right). \tag{5.140}$$

The difference between the two orderings is given by

$$\hat{\mathbf{p}}\cdot\hat{\mathbf{r}}\,\frac{1}{r} = \frac{1}{r}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} - \frac{2i\hbar}{r}. \tag{5.141}$$

Hence, the radial momentum is given by

$$\hat{p}_r = \frac{1}{r}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} - \frac{i\hbar}{r}. \tag{5.142}$$

To obtain the radial part of the Hamiltonian, we have to square this expression. This gives rise to additional commutators:

$$\begin{aligned}\hat{p}_r^2 &= \left(\frac{1}{r}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} - \frac{i\hbar}{r}\right)^2 \\ &= \frac{1}{r^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 + \frac{1}{r}\left[\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right), \frac{1}{r}\right]\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right) - \frac{2i\hbar}{r^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right) - \frac{i\hbar}{r}\left[\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right), \frac{1}{r}\right] - \frac{\hbar^2}{r^2} \\ &= \frac{1}{r^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 - \frac{i\hbar}{r^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right).\end{aligned} \tag{5.143}$$

We used here the commutator expression

$$\left[\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right), \frac{1}{r}\right] = \frac{i\hbar}{r}. \tag{5.144}$$

---

[11] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford: Clarendon Press, 1930.

The resulting radial kinetic energy is given by

$$\hat{\mathscr{H}}_{rad} = \frac{1}{2m}\hat{p}_r^2 = \frac{1}{2mr^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 - \frac{i\hbar}{2mr^2}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right). \tag{5.145}$$

When this part is summed with the angular part, as given in Eq. (5.138), we indeed recover the full kinetic energy operator. In other words, the kinetic energy of a free particle can be reexpressed in spherical components as

$$\hat{\mathscr{H}} = \frac{1}{2mr^2}\hat{L}^2 + \frac{1}{2m}\hat{p}_r^2. \tag{5.146}$$

For a particle confined to moving on a spherical surface with constant radius R, no radial motion is possible and the Hamiltonian reduces to its angular part. That is,

$$\hat{\mathscr{H}} = \frac{1}{2mR^2}\hat{L}^2. \tag{5.147}$$

Hence, as in the cyclic case, the Hamiltonian of the spherical problem is proportional to the square of the angular momentum. As a consequence, the Hamiltonian is seen to commute with the Casimir operator:

$$\left[\hat{\mathscr{H}}, \hat{L}^2\right] = 0, \tag{5.148}$$

and since $\left[\hat{L}^2, \hat{L}_i\right] = 0$ $(i = 1,2,3)$, the Hamiltonian also commutes with the components of the angular momentum operator $\hat{L}$:

$$\left[\hat{\mathscr{H}}, \hat{L}_i\right] = 0. \tag{5.149}$$

The quantum states of this problem will thus fully coincide with the symmetry states of the spherical symmetry group.

5.7.2 The flooded planet model and Buckminsterfullerene

Eigenstates of the spherical Hamiltonian are characterized by eigenvalues of the Casimir operator, which thus recognizes the spherical shells. However, although all the components of the same shell are recognized by the Casimir operator as belonging to the same eigenvalue, the recognition of the individual components in the shell is the task of the component operators $\hat{L}_x$, $\hat{L}_y$, or $\hat{L}_z$.

Now we turn to the particle-on-a-sphere model, which is also called the *hollow sphere* or *flooded planet model*. The eigenstates of the Laplacian on the surface are spherical harmonics that describe a pattern of waves on a sphere, such as on a planet covered with a giant sea. Because the angular Hamiltonian is proportional to the Casimir operator $\hat{L}^2$, its eigenstates correspond to the angular momentum eigenstates, and are characterized by j and m_j labels. In the SO(3) group, angular momentum eigenstates can be integer or half-integer. For a particle on a sphere, only integer values of j are allowed, for the same reason as in cyclic symmetry; that is, functions should be single-valued in periodic angular coordinates. Usually, the integer values of angular momentum are denoted as l-levels, with m_l components. It is convenient to use the symbolic ket representation of Dirac to rewrite Eqs. (5.126) and (5.127) as

$$\hat{L}^2\,|l,m_l\rangle = l(l+1)\,\hbar^2\,|l,m_l\rangle; \tag{5.150}$$

$$\hat{L}_z\,|l,m_l\rangle = m_l\hbar\,|l,m_l\rangle. \tag{5.151}$$

The energies of the eigenstates are then given by

$$E_l = \frac{l\,(l+1)\,\hbar^2}{2mR^2}. \tag{5.152}$$

Traditionally, the spherical shells are denoted by the labels s, p, d, and f for $l = 0$, 1, 2, and 3, respectively, as a tribute to the early designations of atomic line spectra. From $l = 5$ onward, labels continue alphabetically starting from g. The magic numbers corresponding to the closing of an l-shell are easily calculated:

$$2\sum_{l=0}^{n}(2l+1) = 2\frac{(2n+1+1)\,(n+1)}{2} = 2\,(n+1)^2\,. \tag{5.153}$$

The magic numbers thus correspond to double squares: 2, 8, 18, 32, 50, 72, ... These numbers correspond to the length of the periods in Mendeleev's table, but there is much more under the surface to be discovered here. The electron counts 8 and 18 occupy a prominent place in chemistry, as the octet rule for first row elements and the eighteen-electron rule for transition metal elements. Both rules are examples of the famous Langmuir noble gas rule, which is presented in Theorem 5.3.

Theorem 5.3 (Langmuir noble gas rule): Through chemical bonds, an atomic element attempts to achieve the outer shell configuration of the noble gas in the same row of the periodic table. ■

Application of the magic numbers has often been proposed as a stability criterion for globular molecules. In analogy to aromatic ring systems obeying Hückel's rule, Hirsch has claimed that globular molecules with the spherical electron counts should be superaromats, as three-dimensional analogues of aromatic rings.[12]

Definition 5.5 (Hirsch rule): A molecule with a globular shape is aromatic when the number of its π-electrons equals $2\,(n+1)^2$, where n is zero or any positive integer. ■

The trouble with this rule is that actually very few globular molecules exist that follow this rule. An analysis of the electronic structure of the C_{60} molecule is very revealing in this respect.[13]

Buckminsterfullerene

When Buckminsterfullerene, *alias* the carbon buckyball C_{60}, was discovered,[14] it was claimed that the π-electronic structure in this molecule would constitute a superaromat, ignoring the simple fact that 60 is not included in the magic number sequence. In fact, when the structure of C_{60} became available (Figure 5.3), it was clear that the bond lengths are alternating, with short bonds (1.40 Å) adjacent to two hexagons, and longer ones (1.46 Å) for the pentagons. To understand what is

---

[12] A. Hirsch, Z. Chen, and H. Jiao. "Spherical Aromaticity in I_h Symmetrical Fullerenes: The $2(N+1)^2$ Rule." *Angewandte Chemie International Edition* 39 (2000), pp. 3915–3917.

[13] A. Ceulemans, S. Compernolle, and E. Lijnen. "Hiatus in the Spherical Shell Model of Fullerenes." *Physical Chemistry Chemical Physics* 6 (2004), pp. 238–241.

[14] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley. "C_{60}-Buckminsterfullerene." *Nature* 318 (1985), pp. 162–163.

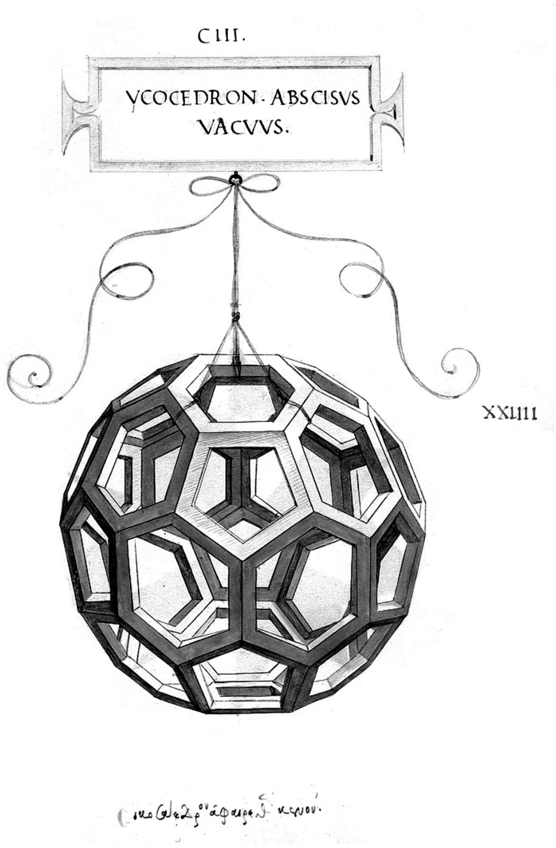


FIGURE 5.3 The so-called *Ycocedron Abscisus Vacuus* from Leonardo Da Vinci is a truncated icosahedron, which is identical in structure to the C_{60} carbon buckyball. From *De Divina Proportione* by Luca Pacioli (Milan, 1509).

going on, we need to examine the orbital sequence of the π-orbitals of C_{60}, as can already be obtained by a low-level quantum mechanical calculation (Figure 5.4).[15] It is clear from the diagram that at lower energies the spherical shell structure clearly imposes itself with consecutive s, p, d, f, and g shells. These accommodate fifty electrons, in agreement with the magic number rule. The next shell, with $l = 5$, however, shows a large splitting; moreover, a component of the $l = 6$ unoccupied shell intrudes into the manifold of the $l = 5$ shell. We are confronted here with a phenomenon that is both disturbing and intriguing at the same time: *spherical symmetry is shattered.* The shell structure is disrupted, but what is the message hidden in this sequence?

The deeper lying orbitals are more bonding, which implies they undergo few sign changes and vary smoothly when going around the sphere. They thus show a close match with spherical waves that are not susceptible to the details of the atomic mesh as they experience a continuous sphere. Orbitals that are less bonding show a more fragmented nodal pattern, and thus probe in more detail the actual atomic structure of the cage. They, indeed, experience that the carbon cage is not a perfect sphere. To cover a sphere with a hexagonal lattice, it is necessary to introduce defects. In the case of C_{60}, the defects are the twelve pentagons, which remind us that the

---

[15] M. R. Savina, L. L. Lohr, and A. H. Francis. “A Particle-on-a-Sphere Model for C_{60}.” *Chemical Physics Letters* 205 (1993), pp. 200–206.

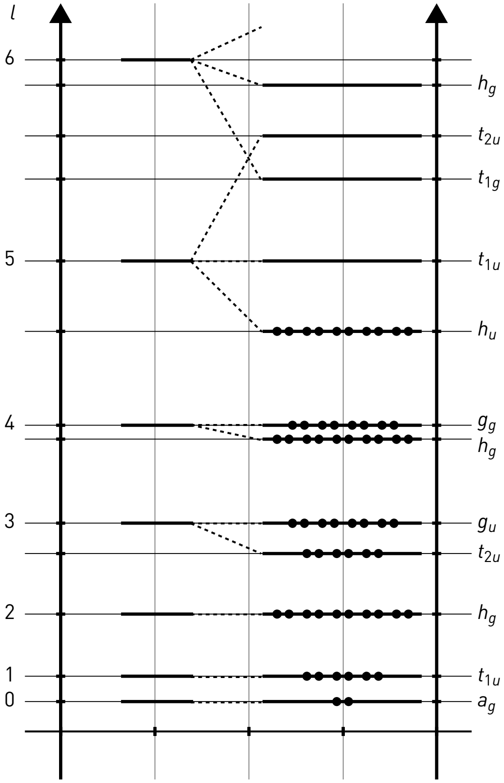


FIGURE 5.4 Orbital sequence of the π-orbitals of C_{60}, adapted from M. R. Savina, L. L. Lohr, and A. H. Francis. "A Particle-on-a-Sphere Model for C_{60}." *Chemical Physics Letters* 205 (1993), pp. 200–206. On the left are the parent spherical levels of the free particle model. On the right are the actual π-orbitals of C_{60}. They are designated by labels for icosahedral symmetry. See A. Ceulemans. *Group Theory Applied to Chemistry*: *Theoretical Chemistry and Computational Modelling.* Dordrecht: Springer, 2013. The HOMO is the h_u-orbital occupied by ten electrons. The diagram displays the breaking of the spherical levels when the symmetry is lowered from SO(3) to icosahedral.

buckyball is obtained by truncating the icosahedron. These defects act as scattering centers that interfere with the spherical waves. The effect becomes clear when the wavelength of the spherical waves becomes of the order of the distance between the scattering defects, and this is precisely what happens near the frontier orbitals. The result is a breaking of the spherical shell and a deviation from the magic shell structure.

Spherical symmetry is thus shattered and shells are fragmented, but the fragments seem to reassemble in a novel type of ordering. In the frontier orbital region, a new ordering principle emerges. As shown by González and colleagues,[16] a continuum model can be applied based on a flat graphite lattice that is curved by insertion of the pentagons. This curving can be mimicked by effective gauge fields emanating from the center of the sphere. At higher energies, into the virtual orbital space toward the top of the π-band, the spectrum can be understood as resulting from an electron moving in the magnetic field of a monopole. For our purposes, we retain that shattering of symmetry may announce the advent of a paradigm that is different from the spherical continuum at the bottom of the band and that takes into account the pentagonal defects in the lattice.

---

[16] J. Gonzalez, F. Guinea, and M. A. H. Vozmediano. "The Electronic Spectrum of Fullerenes from the Dirac Equation." *Nuclear Physics B* 406 (1993), pp. 771–794.

5.8 EPILOGUE

The mathematical treatment of spherical symmetry has revealed that angular momentum is quantized. Interestingly, the quantum numbers representing spherical symmetry can be integer or half-integer. For the current physical system of a particle evolving on a sphere, the half-integer solutions had to be rejected because they lead to double-valued wave functions. The rejection is only temporary. Half-integer solutions will show up again when we will look deeper into the nature of the electron and discover its spin structure in Chapter 8.

6 Scholium I

Noether's theorem towers like an intellectual Mount Everest, the grand peak standing bright and clear over an impressive mountain range of powerful ideas.
–Dwight E. Neuenschwander (2011)[1]

David Hilbert (1862–1943) was furious! He had been sitting in silence for quite some time, overhearing the conversation among his colleagues, but he could not hold his temper any longer. Slapping his fist onto the table, he shouted, "I do not see that the sex of the candidate is an argument against her admission as a *Privatdozent*! After all, *the [faculty] senate is not a bathhouse*!"[2]

Hilbert was professor at the University of Göttingen and he was considered the most influential mathematician of the early twentieth century. Yet, for the last three years he had been fighting a losing battle in a vain attempt to obtain a permanent faculty position for his dear, esteemed colleague Emmy Noether (1882–1935) (Figure 6.1). Try as he might, his requests were met with horror and disapproval.

"What will our soldiers think when they return to the university and find that they are required to learn at the feet of a woman?"[3] one professor complained. The problem, put bluntly, was that Noether lacked a Y chromosome, and women

---

[1] D. E. Neuenschwander. *Emmy Noether's Wonderful Theorem*. Baltimore: Johns Hopkins University Press, 2011, p. xi.

[2] Ibid., p. 7.

[3] Ibid.

FIGURE 6.1 One of the few photographs of Emmy Noether (1882–1935).

during the early 1900s were excluded from holding faculty positions at German universities.[4]

Sadly enough, this was not the first (and certainly not the last) of obstacles that Noether met on her way to academic recognition. A few year earlier, for example, at the age of eighteen, Noether had qualified as a language teacher, but she wished to follow in her father's footsteps; he was an established mathematics professor at the University of Erlangen. After rebelling against her family—girls, after all, were not supposed to study mathematics—Noether was forced to gain personal permission from every professor to enroll in their courses because women were not generally admitted in class. In the end, Noether was allowed to audit classes, but she could not sit for examination. So, she had to pass her examination at a nearby university in Nuremberg. This was in 1903.

Five years later, Noether completed her PhD in mathematics at the University of Göttingen, *summa cum laude*, under the supervision of algebraist Paul A. Gordan (1837–1912). Gordan was a close friend of Emmy Noether's father, Max Noether, and he was known as "the king of invariant theory." Before long, Emmy Noether had become a first-class expert in mathematical invariants. We've encountered *invariants* before; they are deeply intertwined with symmetry. In the words of Dave Goldberg:

> Invariants are the counterpoint to symmetries. While a symmetry describes the sort of transformations that you can apply to a system without changing it, an invariant is the thing itself that is unaltered.[5]

For example, when you rotate a circle around its midpoint, its shape (and radius) remains invariant. More importantly many *laws in physics* are said to be invariant under certain symmetry transformations, in the sense that their functional form does not change under those transformations. We saw examples of this in the last chapter.

For the next eight years, Noether worked, unpaid, at the University of Erlangen, occasionally lecturing for her father.

---

[4] This attitude stood in stark contract with the more liberal universities of France, England, and Italy.

[5] D. Goldberg. *The Universe in the Rearview Mirror. How Hidden Symmetries Shape Reality*. New York: Dutton, 2013. The quote is from Goldberg's blog.

Not everyone was as narrow-minded. The great Albert Einstein, for example, was well aware of Noether's brilliance, and wrote in a letter to Felix Klein, dated December 27, 1918: "On receiving the new work from Fräulein Noether, I again find it a great injustice that she cannot lecture officially. I would be very much in favour of taking energetic steps in the ministry [to overturn this rule]."[6] And in a letter to Hilbert, dated May 24 of that same year, Einstein extorted: "It would not have done the Old Guard at Göttingen any harm, had they picked up a thing or two from her. She certainly knows what she is doing."[7]

But who was this woman—so highly esteemed by some of the greatest mathematicians of her time (such as Hilbert, Klein, and Einstein, but also Hermann Weyl and Herman Minkowski), and yet so despised by the other academicians? Perhaps you never heard of Emmy Noether before?

In this *scholium*, we pause to contemplate Noether's theorem, which explains why group theory is so particularly effective to treat quantum chemical problems.

Chapter Outline

We are about to confront the two great central field potentials of quantum physics: the *harmonic oscillator* with SU(3) symmetry (Chapter 7) and the *Coulomb attraction hole* with SO(4) symmetry (Chapter 9). The purpose of this chapter is to prepare our approach to these problems from the point of view of symmetry.

The relevance of the group concept for physics and chemistry has its origin in the fundamental connection between the physical properties of a system and its symmetry transformations. That is to say, the complete *set* of symmetry transformations of an object or a law, along with the *operation* of combining two such transformations by performing one after the other, forms the *symmetry group* of the Hamiltonian.

To see this more clearly, consider the idea of mirror symmetry. Scientists long believed that the laws of physics were invariant under mirror reflections. Whether one observes Nature through a mirror or looks at it directly, physics still works the same (Figure 6.2). Now, the two possible transformations (i.e., the identity and the reflection) form a symmetry group isomorphic to the group $(\mathcal{C}_2, \times)$ of order 2, with $\mathcal{C}_2 = \{1, -1\}$.

In this chapter, we will study the proper place of symmetry operations in the framework of quantum mechanics (§6.1). We will be interested, in particular, in symmetry operators generating Lie algebras (§6.2). Lie generators play a central role in quantum mechanics because they can be represented by Hermitian operators, which in turn correspond to physical observables. As will be shown in §6.3, this leads to conserved quantities (or *constants of the motion*), and this observation sets the scene for a treatment of the profound relation that exists between symmetries and conservation laws, as first described by Noether's theorem (§6.3.1).

In a final section, §6.4, we will delve even deeper into the variegated anatomy of Lie algebras to reveal their inner workings. One way of uncovering the full power of a Lie algebra is by introducing its *Cartan subalgebra* and associated *Weyl diagram* by way of the so-called *Cartan-Weyl method*. The importance of this method cannot be stressed enough; time and again, the Cartan-Weyl method will reveal itself as a crucial

---

[6] Neuenschwander. *Emmy Noether's Wonderful Theorem*, p. 8.

[7] Ibid., p. 8.

FIGURE 6.2 Alice through the looking glass. Scientists long thought the laws of science would stay the same if you stepped in a mirror world, like Alice. We now know that the symmetry between left and right is sometimes violated by Nature. This is referred to as *parity violation.*

tool in our analysis of Lie groups and their corresponding Lie algebras. This will be exemplified by a brief consideration of the *angular momentum algebra* in §6.4.2.

6.1 SYMMETRY IN QUANTUM MECHANICS

Quantum mechanics is based on wave functions that describe the state of a physical system. An important distinction has to be made, however, between the transformations of space–time coordinates used to describe a system in physical space and the corresponding transformations of the wave function that describes the state of the system in Hilbert space (§6.1.1).

With this distinction in place, we can study the nature of these quantum mechanical transformations (§6.1.2) and analyze how observables transform under their action (§6.1.3). Central among the observables in quantum mechanics is the total energy E of the system. Because this observable is represented by the Hamiltonian operator $\hat{\mathscr{H}}$, it will prove interesting to examine how $\hat{\mathscr{H}}$ transforms under a given symmetry transformation (§6.1.4). A proof will be given that the symmetry transformations of a quantum mechanical system indeed form a group: the symmetry group of the Hamiltonian (§6.1.5). This consideration will lead to one of the most important results of this chapter: *symmetry gives rise to degeneracy*. The greater the degeneracies observed in a system, the greater the symmetry underlying that system (§6.1.6). The exploration of this idea will form a recurrent team in the chapters to follow.

6.1.1 State vector transformations

Let $\mathscr{S}$ denote a particular quantum system. To each *coordinate transformation T* that induces a linear transformation of the space–time coordinates of $\mathscr{S}$, there corresponds a *function operator* $\hat{T}$ that operates on the wave functions ψ of the system and that transforms them into new functions ψ'.

It will be convenient to introduce Dirac's *bra-ket* notation. The wave function ψ is then written as a *ket* $|\psi\rangle$, and the action of $\hat{T}$ on the ket $\left|\psi\left(x,y,z,t\right)\right\rangle = |\psi\left(\mathbf{r},t\right)\rangle$ is given by the following equation:

$$\hat{T}\,|\psi\left(\mathbf{r},t\right)\rangle = \left|\psi'\left(\mathbf{r},t\right)\right\rangle. \tag{6.1}$$

The operator $\hat{T}$ is said to *represent* the transformation T and induces a "rotation" of the wave function ψ in the abstract Hilbert space $\mathcal{H}$ of $\mathscr{S}$.

To determine the functional form of the new (transformed) wave function $\left|\psi'\right\rangle$, use can be made of the fact that $\left|\psi'\right\rangle$ is a function and $\hat{T}$ is a function operator. We can, therefore, apply the derivation of Chapter 3, §3.5, for the transformation of coordinate functions to the wavefunction, yielding

$$\hat{T}\,|\psi\left(\mathbf{r}\right)\rangle = \left|\psi'\left(\mathbf{r}\right)\right\rangle = \left|\psi\left(T^{-1}\mathbf{r}\right)\right\rangle. \tag{6.2}$$

Because the wave functions represent a set of vectors, we can set up a second set of vectors that form a dual space. The elements of the latter are denoted as bra functions $\langle\psi\left(\mathbf{r},t\right)|$. The scalar product of a bra and a ket is then written as a bracket, $\langle\psi\,|\,\psi\rangle$, which is a scalar number. Along this line, the symmetry transformation in the dual bra space is as

$$\left\langle\hat{T}\psi\left(\mathbf{r},t\right)\right| = \left\langle\psi'\left(\mathbf{r},t\right)\right| = \langle\psi|\,\hat{T}^{\dagger}, \tag{6.3}$$

with $\hat{T}^{\dagger}$ denoting the *adjoint* (or *Hermitian conjugate*) of $\hat{T}$ (see Appendix D).

6.1.2 Unitarity of symmetry operators

Let us inquire into the nature of the operator $\hat{T}$, which was introduced in §6.1.1. Suppose a system is in a state $|\phi\rangle$. Then, the probability that a measurement of an operator will yield an eigenvalue that corresponds to the eigenfunction $|\psi\rangle$ of this operator will be $|\langle\phi\mid\psi\rangle|^2$. It is required that this probability remains invariant under a symmetry operation $\hat{T}$

$$|\langle\phi\mid\psi\rangle|^2 = \left|\left\langle\phi'\mid\psi'\right\rangle\right|^2. \tag{6.4}$$

One way to satisfy this equation is by postulating that the scalar product $\langle\phi\mid\psi\rangle$ remains invariant:

$$\langle\phi\mid\psi\rangle = \left\langle\phi'\mid\psi'\right\rangle = \langle\phi\mid\hat{T}^{\dagger}\,\hat{T}\mid\psi\rangle, \tag{6.5}$$

where Eqs. (6.1) and (6.3) have been used. For this equality to hold,

$$\hat{T}^{\dagger}\,\hat{T} = \hat{T}\hat{T}^{\dagger} = \hat{I}, \tag{6.6}$$

with $\hat{I}$ being the identity operator. This means that the adjoint of the symmetry operator $\hat{T}$ coincides with the inverse operator:

$$\hat{T}^{\dagger} = \hat{T}^{-1}. \tag{6.7}$$

Operators with this property are said to be *unitary*. Generalizing this finding leads to Wigner's famous theorem[8]:

---

[8] This theorem originated in the writings of Eugene Wigner (1902–1995) in 1931. See E. P. Wigner. *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra*. New York: Academic

Theorem 6.1 (Wigner's theorem): Every operator $\hat{T}$ that corresponds to a symmetry transformation T must be unitary; that is, $\hat{T}^\dagger = \hat{T}^{-1}$. ▪

6.1.3 Transformation of observables

Next, let us investigate how observables change under a transformation $\hat{T}$. The expectation value $\left\langle \hat{A} \right\rangle$ of an observable A (corresponding to the operator $\hat{A}$), with respect to the original state $|\psi\rangle$ is given by

$$\left\langle \hat{A} \right\rangle = \left\langle \psi \left| \hat{A} \right| \psi \right\rangle. \tag{6.8}$$

This expectation value must remain invariant during a symmetry transformation $\hat{T}$. That is, the expectation value of A with respect to the original state $|\psi\rangle$ should be the same as the expectation value of the transformed observable A' with respect to the transformed state $|\psi'\rangle$. Denoting this transformed expectation value by $\left\langle \hat{A}' \right\rangle = \left\langle \psi' \left| \hat{A}' \right| \psi' \right\rangle$, we get

$$\left\langle \psi \left| \hat{A} \right| \psi \right\rangle = \left\langle \psi' \left| \hat{A}' \right| \psi' \right\rangle = \left\langle \hat{T}\psi \left| \hat{A}' \right| \hat{T}\psi \right\rangle = \left\langle \psi \left| \hat{T}^\dagger \hat{A}' \hat{T} \right| \psi \right\rangle. \tag{6.9}$$

This leads to the following equality:

$$\hat{A} = \hat{T}^\dagger \hat{A}' \hat{T}, \tag{6.10}$$

or, using Eq. (6.7),

$$\hat{A}' = \hat{T} \hat{A} \hat{T}^\dagger, \tag{6.11}$$

relating the untransformed and transformed observables $\hat{A}$ and $\hat{A}'$ via the unitary operator $\hat{T}$ that was used to describe the rotation of the state ket $|\psi\rangle$.

If the observable $\hat{A}$ is *invariant* under $\hat{T}$ (i.e., $\hat{T}\hat{A}\hat{T}^\dagger = \hat{A}' = \hat{A}$), then the transformed observable $\hat{A}'$ equals $\hat{A}$, and the previous equation reduces to

$$\hat{A}' = \hat{T}\hat{A}\hat{T}^\dagger = \hat{A}, \tag{6.12}$$

which implies that $\hat{T}\hat{A} = \hat{A}\hat{T}$, and thus

$$\left[\hat{A}, \hat{T}\right] = \hat{A}\hat{T} - \hat{T}\hat{A} = 0. \tag{6.13}$$

Definition 6.1 (Invariance of observables): An observable $\hat{A}$ is said to be *invariant* under a symmetry operator $\hat{T}$ when it commutes with the operator $\hat{T}$: $\left[\hat{A}, \hat{T}\right] = 0$. The observable $\hat{A}$ is then referred to as a *scalar* with respect to the symmetry operation. ▪

In quantum mechanics operators, associated with observables, obey the condition of *Hermiticity*:

$$\left\langle \phi \,|\, \hat{A}\psi \right\rangle = \left\langle \hat{A}\phi \,|\, \psi \right\rangle. \tag{6.14}$$

This implies that quantum mechanical operators are *self-adjoint*:

$$\hat{A}^\dagger = \hat{A}. \tag{6.15}$$

Press, 1959. It must be remarked that Theorem 6.1 is a simplified version of Wigner's; it does not mention the existence of *antiunitary* symmetry operators, which are used to describe *time-reversal symmetry*.

The reason for this condition is a physical requirement. Physical observables must be *real* quantities, implying the product $\left\langle \psi|\hat{A}|\psi\right\rangle$ must be a *real scalar*. It can be easily shown that this is the case for Hermitian operators. For $\hat{A}$ Hermitian, we have

$$\left\langle \psi|\hat{A}\psi\right\rangle = \left\langle \hat{A}\psi|\psi\right\rangle = \left\langle \psi|\hat{A}\psi\right\rangle^{*}. \tag{6.16}$$

6.1.4 Symmetry transformations of the Hamiltonian

Let $\hat{\mathscr{H}}$ represent the Hamiltonian of a particular quantum system $\mathscr{S}$, with $\mathcal{H}$ the Hilbert space of eigenvectors $|\psi\rangle$ for which the time-independent Schrödinger equation applies:

$$\hat{\mathscr{H}}|\psi\rangle = E|\psi\rangle. \tag{6.17}$$

Acting with $\hat{T}$ on both sides of the Schrödinger equation yields

$$\hat{T}\hat{\mathscr{H}}|\psi\rangle = \hat{T}E|\psi\rangle, \tag{6.18}$$

or

$$\hat{T}\hat{\mathscr{H}}\hat{T}^{-1}\hat{T}|\psi\rangle = E\hat{T}|\psi\rangle. \tag{6.19}$$

Furthermore, suppose $\hat{\mathscr{H}}$ is *invariant* under the transformation $\hat{T}$—that is,

$$\hat{T}\hat{\mathscr{H}}\hat{T}^{-1} = \hat{\mathscr{H}}' = \hat{\mathscr{H}}. \tag{6.20}$$

In this case, $\hat{T}$ is called a *symmetry transformation* of $\hat{\mathscr{H}}$. Alternatively, Eq. (6.20) can be written as

$$\hat{T}\hat{\mathscr{H}} = \hat{\mathscr{H}}\hat{T} \tag{6.21}$$

or

$$\left[\hat{\mathscr{H}}, \hat{T}\right] = 0. \tag{6.22}$$

Definition 6.2 (Invariance of the Hamiltonian): Let $\hat{\mathscr{H}}$ denote the Hamiltonian of a quantum system $\mathscr{S}$ that is invariant under the action of a *symmetry operator* $\hat{T}$. Then, $\hat{\mathscr{H}}$ *commutes* with the operator $\hat{T}$: $\left[\hat{\mathscr{H}}, \hat{T}\right] = 0$. ■

6.1.5 Symmetry group of the Hamiltonian

By combining all the symmetry transformations of the Hamiltonian we can construct the full symmetry group.

Theorem 6.2 Let $\{T\}$ be the set of coordinate transformations and $\{\hat{T}\}$ be the set of corresponding function operators under which the Hamiltonian $\hat{\mathscr{H}}$ remains invariant. Then, $\{T\}$ forms a *group* $\mathcal{G}$, with $\{\hat{T}\}$ a *representation* of that group. ■

Proof that $\{T\}$ *forms a group*: It is evident that the sequential execution of two symmetry transformations, T_2T_1, is also a symmetry transformation T_3. That is, if $T_1\mathbf{r} = \mathbf{r}'$ and $T_2\mathbf{r}' = \mathbf{r}''$, then $T_3\mathbf{r} = \mathbf{r}''$. The associative law also holds since the execution of T_1 and then T_3T_2 is the same as first executing T_2T_1 and then T_3. The identity transformation I leaves the spatial points untouched: $I\mathbf{r} = \mathbf{r}$. Finally, for each transformation T, there is an inverse transformation T^{-1} that undoes the first transformation; in other words, if $T\mathbf{r} = \mathbf{r}'$, then $T^{-1}\mathbf{r}' = \mathbf{r}$. This proves that the set $\{T\}$ forms a group $\mathcal{G}$. ■

Proof that $\{\hat{T}\}$ *forms a representation of* $\mathcal{G}$: Now, let T_1 and T_2 denote two coordinate operators and suppose that

$$T_2 T_1 = T_3. \tag{6.23}$$

Following Eq. (6.2), we describe the action of the corresponding function operators $\hat{T}_1$ and $\hat{T}_2$ as follows:

$$\hat{T}_1 \left|\psi(\mathbf{r})\right\rangle = \left|\psi'(\mathbf{r})\right\rangle = \left|\psi\left(T_1^{-1}\mathbf{r}\right)\right\rangle; \tag{6.24}$$

$$\hat{T}_2 \left|\psi'(\mathbf{r})\right\rangle = \left|\psi''(\mathbf{r})\right\rangle = \left|\psi'\left(T_2^{-1}\mathbf{r}\right)\right\rangle. \tag{6.25}$$

Substituting $T_2^{-1}\mathbf{r}$ for $\mathbf{r}$ in the definition of $\left|\psi'\right\rangle$ yields

$$\left|\psi'\left(T_2^{-1}\mathbf{r}\right)\right\rangle = \left|\psi\left(T_1^{-1}\left(T_2^{-1}\mathbf{r}\right)\right)\right\rangle = \left|\psi\left(T_1^{-1}T_2^{-1}\mathbf{r}\right)\right\rangle = \left|\psi\left((T_2T_1)^{-1}\mathbf{r}\right)\right\rangle. \tag{6.26}$$

Substitution of this result in Eq. (6.25) finally gives

$$\hat{T}_2\hat{T}_1 \left|\psi(\mathbf{r})\right\rangle = \left|\psi\left((T_2T_1)^{-1}\mathbf{r}\right)\right\rangle = \left|\psi\left(T_3^{-1}\mathbf{r}\right)\right\rangle = \hat{T}_3 \left|\psi(\mathbf{r})\right\rangle, \tag{6.27}$$

or simply

$$\hat{T}_2\hat{T}_1 = \hat{T}_3. \tag{6.28}$$

This shows that the operators $\hat{T}$ are multiplied *in the same order* as the transformations T. Also, the product of two operators $\hat{T}_2\hat{T}_1$ is seen to yield another operator $\hat{T}_3$. This corresponds to the requirement of *closure*.

For T_2 equal to T_1^{-1} in Eq. (6.27), we get

$$\hat{T}_2\hat{T}_1 = \hat{I}, \tag{6.29}$$

with $\hat{I}$ as the *identity operator*. It follows that the operator associated with T_1^{-1} is the *inverse* of $\hat{T}_1$.

This verifies all group postulates and proves that the symmetry operators $\{\hat{T}\}$ form a group. Because each element T of $\mathcal{G}$ is assigned an operator $\hat{T}$ operating in the Hilbert space $\mathcal{H}$ of eigenvectors $\psi(\mathbf{r})$, it is said that the operators $\{\hat{T}\}$ form a *representation* of the group $\mathcal{G}$. ■

6.1.6 Symmetry and degeneracy

Whenever a symmetry of the Hamiltonian has been revealed, *degeneracies* are seen to arise, and the more symmetry, the higher the degeneracy. To see this more clearly, let us rewrite Eq. (6.19) in light of Eq. (6.20) to yield

$$\hat{\mathscr{H}}\hat{T}\left|\psi\right\rangle = E\hat{T}\left|\psi\right\rangle. \tag{6.30}$$

Rewriting the new state vector $\hat{T}\left|\psi\right\rangle$ as $\left|\psi'\right\rangle$ according to Eq. (6.1), we obtain

$$\hat{\mathscr{H}}\left|\psi'\right\rangle = E\left|\psi'\right\rangle. \tag{6.31}$$

Two possibilities arise at this point:

1. *Nondegenerate case*: There is only one eigenfunction with energy E. In this case, the transformed wave function $\left|\psi'\right\rangle$ must be the same as the original wave function $\left|\psi\right\rangle$, except possibly for a unimodular phase factor: $\left|\psi'\right\rangle = e^{i\kappa}\left|\psi\right\rangle$, with $\kappa \in \mathbb{R}$ a real number.

2. *Degenerate case*: There is a set of eigenfunctions that share the eigenenergy E. This set can be denoted as $\{|\psi_i\rangle \mid i = 1, 2, \ldots, n\}$, with n the set cardinality. This set is also referred to as the *eigenspace* of E, and the elements of the set are said to be *degenerate* eigenkets. Any transformed wave function $\left|\psi_i'\right\rangle$ also belongs to this set and must therefore coincide with a linear combination of the elements of the eigenspace of E:

$$\hat{T}\,|\psi_i\rangle = \sum_{j=1}^{n} c_{ij}\left|\psi_j\right\rangle. \tag{6.32}$$

   This equation illustrates the important relationship between symmetry and degeneracy; the symmetry operators are able to connect the elements of the eigenspace, something the Hamiltonian is incapable of doing. In fact, the mere existence of symmetry operators is the fundamental reason for the existence of degeneracy.

The symmetry transformations $\hat{T}$ of the Hamiltonian $\hat{\mathscr{H}}$ can thus be used to relate the different eigenkets of one energy level (E) with one another. This allows us to discuss the degree of degeneracy of that particular energy level—an important line of thought that will be developed further in §6.4.1.

6.2 LIE GROUPS AND LIE ALGEBRAS

Discrete transformations give rise to *discrete symmetry groups*. According to the standard model of particle physics, for instance, the laws of physics are invariant under charge conjugation transformations, parity transformations, and time-reversal transformations (Figure 6.3). The corresponding C, P, and T symmetries each consist of two symmetry operators, which could be labeled by a discrete parameter that takes

FIGURE 6.3 The characters of Tweedledee and Tweedledum are symmetric reflections of one another. They have been referred to as an enantiomeric pair, mirror twins, or a matter particle and its corresponding antimatter particle.

only two values—say, 0 and 1; $\hat{T}(0) = \hat{I}$ would denote the identity operator, and $\hat{T}(1)$ the $\hat{C}$, $\hat{P}$, or $\hat{T}$ operator.

In many cases, however, the parameters vary over a continuous interval, and *continuous symmetry groups* arise, as we saw in Chapters 4 and 5. These groups are called *Lie groups* and because the group elements $\hat{T}(r_1, r_2, \ldots, r_n)$ depend on the parameters $\boldsymbol{\rho} = \{r_1, r_2, \ldots, r_n\}$ in a differentiable manner, it is possible to study *infinitesimal* symmetry operations; these are the *generators* of a symmetry because all group elements $\hat{T}(r_1, r_2, \ldots, r_n)$ can be obtained from this small set of generators (§6.2.1).

We have seen that any two elements of a Lie group can be combined to form a third element. In a similar vein, two infinitesimal transformations can be combined to yield a third transformation, as demonstrated in §5.4. Interestingly, this operation no longer has the familiar properties of a Lie group multiplication. Rather, a new algebraic structure is seen to arise, which is referred to as a *Lie algebra* (§6.2.2).

6.2.1 Lie generators

Let $\mathcal{G}$ denote a Lie group, the elements of which are characterized by a finite set of n real, independent parameters r_i $(i = 1, \ldots, n)$ that vary continuously over a particular interval. Denoting the elements of the group by $\hat{T}(r_1, r_2, \ldots, r_n)$, or $\hat{T}(\boldsymbol{\rho})$ for short, with $\boldsymbol{\rho} = \{r_1, r_2, \ldots, r_n\}$, the group $\mathcal{G}$ is said to be an *n-parameter* group or to be a continuous group of *order n.*

Definition 6.3 (Lie generators of a Lie group): To every unitary symmetry operator $\hat{T}(\boldsymbol{\rho})$ of a Lie group $\mathcal{G}$, which depends analytically on the parameters $\boldsymbol{\rho} = \{r_1, r_2, \ldots, r_n\}$, there is a corresponding set of n Lie generators $\hat{X}_i$ according to the equation

$$\hat{T}(\boldsymbol{\rho}) = \exp\left(\sum_{i=1}^{n} r_i \hat{X}_i\right), \tag{6.33}$$

where

$$\hat{X}_i = \left.\frac{\partial \hat{T}(\boldsymbol{\rho})}{\partial r_i}\right|_{r_i=0} \tag{6.34}$$

are the different generators. ■

6.2.2 Lie algebras

The infinitesimal generators of a Lie group $\mathcal{G}$ do not in any way *belong* to the Lie group they generate. Instead, as with the components $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ of the angular momentum operator, they are seen to form a very different kind of structure called the *Lie algebra* $\mathfrak{g}$ of the group $\mathcal{G}$.[9] In general, an n-parameter Lie group gives rise to n generators $\hat{X}_i$, which form an n-dimensional Lie algebra.

Definition 6.4 (Lie algebra): An n-dimensional Lie algebra $\mathfrak{g}$ is a linear vector space over some field $\mathbb{F}$, spanned by n generators $\hat{X}_i$ $(i = 1, \ldots, n)$, that is equipped

---

[9] Lie algebras are usually denoted by small Gothic letters $\mathfrak{a}, \mathfrak{b}, \ldots, \mathfrak{g}, \ldots$. Because, in this book, every Lie algebra will be associated with a Lie group, we will use the spelling of the Lie group. Thus, the Lie algebra of the Lie group SO(2), for example, is denoted $\mathfrak{so}(2)$, as described earlier.

with a multiplication operator (i.e., product rule) $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the *Lie bracket*:

$$\left[\hat{X}_i, \hat{X}_j\right] = \sum_{k=1}^{n} f_{ijk} \hat{X}_k \equiv f_{ijk} \hat{X}_k, \tag{6.35}$$

with f_{ijk} the different *structure constants*. ∎

The Lie bracket (see also Appendix F) is defined as

$$\left[\hat{X}_i, \hat{X}_j\right] := \hat{X}_i \star \hat{X}_j - \hat{X}_j \star \hat{X}_i. \tag{6.36}$$

If the generators are written in matrix form, the product symbol $\star$ denotes matrix multiplication and the Lie bracket corresponds to the *commutator* of $\hat{X}_i$ and $\hat{X}_j$. If the generators are written in operator form, the product symbol $\star$ stands for the consecutive application of $\hat{X}_i$ and $\hat{X}_j$. The Lie bracket should satisfy the following properties:

1. *Bilinearity*: The Lie bracket is bilinear for all scalars λ, μ in $\mathbb{F}$ and all elements $\hat{X}_i$, $\hat{X}_j$, and $\hat{X}_k$ in $\mathfrak{g}$:

$$\begin{aligned} \left[\lambda \hat{X}_i + \mu \hat{X}_j, \hat{X}_k\right] &= \lambda \left[\hat{X}_i, \hat{X}_k\right] + \mu \left[\hat{X}_j, \hat{X}_k\right]; \\ \left[\hat{X}_k, \lambda \hat{X}_i + \mu \hat{X}_j\right] &= \lambda \left[\hat{X}_k, \hat{X}_i\right] + \mu \left[\hat{X}_k, \hat{X}_j\right]. \end{aligned} \tag{6.37}$$

2. *Self-commutation*: The Lie bracket satisfies

$$\left[\hat{X}_i, \hat{X}_i\right] = 0. \tag{6.38}$$

3. *Jacobi identity*: The Lie bracket satisfies the Jacobi identity for all $\hat{X}_i$, $\hat{X}_j$, and $\hat{X}_k$ in $\mathfrak{g}$:

$$\left[\hat{X}_i, \left[\hat{X}_j, \hat{X}_k\right]\right] + \left[\hat{X}_j, \left[\hat{X}_k, \hat{X}_i\right]\right] + \left[\hat{X}_k, \left[\hat{X}_i, \hat{X}_j\right]\right] = 0. \tag{6.39}$$

The first and second properties in Eqs. (6.37) and (6.38) imply that

$$\begin{aligned} 0 = \left[\hat{X}_i + \hat{X}_j, \hat{X}_i + \hat{X}_j\right] &= \left[\hat{X}_i, \hat{X}_i\right] + \left[\hat{X}_i, \hat{X}_j\right] + \left[\hat{X}_j, \hat{X}_i\right] + \left[\hat{X}_j, \hat{X}_j\right] \\ &= \left[\hat{X}_i, \hat{X}_j\right] + \left[\hat{X}_j, \hat{X}_i\right], \end{aligned}$$

or

$$\left[\hat{X}_i, \hat{X}_j\right] = -\left[\hat{X}_j, \hat{X}_i\right]. \tag{6.40}$$

This is known as the *antisymmetry* or *skew-symmetry* property.

The dimension $\dim \mathfrak{g} = n$ of a Lie algebra is defined as the dimension of the vector space $\mathfrak{g}$, spanned by the generators $\hat{X}_i$ $(i = 1, \ldots, n)$. A *basis* $\mathcal{B}$ can then be introduced:

$$\mathcal{B} = \{\hat{X}_i \mid i = 1, 2, \ldots, n\}. \tag{6.41}$$

The structure constants f_{ijk} in Eq. (6.35) contain all the structural information needed to define a Lie algebra uniquely; knowledge of these constants allows the computation of every possible commutator. To see this, let $\hat{X}$ denote a general element from the Lie algebra $\mathfrak{g}$. Since the n generators $\hat{X}_i$ form a *basis* of the Lie algebra $\mathfrak{g}$, every $\hat{X} \in \mathfrak{g}$ can be written as a linear combination of the $\hat{X}_i$'s:

$$\hat{X} = \sum_{i=1}^{n} c_i \hat{X}_i \equiv c_i \hat{X}_i; \qquad c_i \in \mathbb{F}, \tag{6.42}$$

where we have used Einstein's summation convention in the second equality. By Eqs. (6.35) and (6.42), the commutator of any two general elements $\hat{X} = \sum_{i=1}^{n} c_i \hat{X}_i$ and $\hat{X}' = \sum_{j=1}^{n} c_j' \hat{X}_j$ of $\mathfrak{g}$ can be written as

$$\begin{aligned} \left[\hat{X}, \hat{X}'\right] &= \left[\sum_{i=1}^{n} c_i \hat{X}_i, \sum_{j=1}^{n} c_j' \hat{X}_j\right] \\ &= \sum_{i,j=1}^{n} c_i c_j' \left[\hat{X}_i, \hat{X}_j\right] \\ &= \sum_{i,j,k=1}^{n} c_i c_j' f_{ijk} \hat{X}_k. \end{aligned} \tag{6.43}$$

This illustrates that the commutator of *any* two elements of $\mathfrak{g}$ can be determined from a knowledge of the structure constants f_{ijk}.

The skew-symmetric property of the Lie bracket in Eqs. (6.38) and (6.40) can be restated in terms of the structure constants as

$$f_{iik} = 0 \quad \text{or} \quad f_{ijk} = -f_{jik}. \tag{6.44}$$

That is, the structure constants are antisymmetric in their first two indices. The Jacobi identity in Eq. (6.39) becomes

$$\sum_{m=0}^{d} \left(f_{jkl} f_{ilm} + f_{kil} f_{jlm} + f_{ijl} f_{klm}\right) = 0. \tag{6.45}$$

The connection between the n generators $\hat{X}_i$ of a Lie algebra $\mathfrak{g}$ and the group elements $\hat{T}(\boldsymbol{\rho})$ of a Lie group $\mathcal{G}$, as given in Eq. (6.33), is referred to as the *exponential map* and it forms one of the fundamental theorems of Lie's theory.

Definition 6.5 (Exponential map): Each $\hat{T}(\boldsymbol{\rho}) \in \mathcal{G}$ can be expressed in terms of a finite set of n generators $\hat{X}_i$ via the *exponential map* $\exp : \mathfrak{g} \to \mathcal{G}$, given by

$$\hat{T}(\boldsymbol{\rho}) = \exp\left(\sum_{i=1}^{n} r_i \hat{X}_i\right),$$

where the $\hat{X}_i$ are elements of the corresponding Lie algebra $\mathfrak{g}$. ■

We end this section with two definitions that will be needed in §6.4.1:

Definition 6.6 (Abelian Lie algebra): An *Abelian* Lie algebra $\mathfrak{g}$ is a Lie algebra for which all structure constants f_{ijk} are equal to zero. As a result, $\left[\hat{X}, \hat{X}'\right] = 0$ for all elements $\hat{X}$ and $\hat{X}'$ of the Lie algebra. ■

Definition 6.7 (Subalgebra of a Lie algebra): Let $\mathfrak{g}$ represent a Lie algebra of dimension n, spanned by the generators $\hat{X}_i$ $(i = 1, \ldots, n)$, and equipped with the commutator $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Then, $\mathfrak{H}$ is called an m-dimensional *subalgebra* of $\mathfrak{g}$ (with $m \leq n$) if it contains a subset of elements of $\mathfrak{g}$ that themselves form a Lie algebra under the same commutation rule as $\mathfrak{g}$. If $m < n$, $\mathfrak{H}$ is said to be a *proper* subalgebra of $\mathfrak{g}$, denoted $\mathfrak{g} \supset \mathfrak{H}$. ■

6.2.3 Hermiticity and Lie generators

Let $\hat{T}(\delta r_i)$ denote an infinitesimal transformation over δr_i, where r_i is one of the n parameters specifying the group element $\hat{T}$. The operator $\hat{T}(\delta r_i)$ differs infinitesimally from the unit operator $\hat{T}(0) = \hat{I}$:

$$\hat{T}(\delta r_i) = \hat{I} + \delta r_i \hat{X}_i + \mathscr{O}\left(\delta r_i^2\right), \tag{6.46}$$

where $\mathscr{O}\left(\delta r_i^2\right)$ denotes terms of higher order in δr_i. Following Wigner's theorem (Theorem 6.1) $\hat{T}(\delta r_i)$ must be unitary:

$$\hat{T}^{\dagger}(\delta r_i)\,\hat{T}(\delta r_i) = \hat{I}. \tag{6.47}$$

Writing the adjoint of Eq. (6.46) as

$$\hat{T}^{\dagger}(\delta r_i) = \hat{I} + \delta r_i \hat{X}_i^{\dagger}, \tag{6.48}$$

and substituting in Eq. (6.47) gives

$$\left(\hat{I} + \delta r_i \hat{X}_i^{\dagger}\right)\left(\hat{I} + \delta r_i \hat{X}_i\right) = \hat{I}. \tag{6.49}$$

Working out the brackets and keeping everything to first order yields

$$\hat{I} + \delta r_i \left(\hat{X}_i + \hat{X}_i^{\dagger}\right) = \hat{I}, \tag{6.50}$$

and thus

$$\hat{X}_i^{\dagger} = -\hat{X}_i. \tag{6.51}$$

Hence, the generator $\hat{X}_i$ of the unitary transformation $\hat{T}(r_i)$ is seen to be *skew-Hermitian* instead of Hermitian. We can easily turn it into an Hermitian operator, however, by multiplying by the imaginary unit i. Let

$$\hat{Y}_k = i\hat{X}_k. \tag{6.52}$$

The adjoint operation then turns i into $-i$ because it involves complex conjugation between the ket and bra parts. As a result, the new operator can be shown to be Hermitian:

$$\hat{Y}_k^{\dagger} = \left(i\hat{X}_k\right)^{\dagger} = -i\hat{X}_k^{\dagger} = i\hat{X}_k = \hat{Y}_k. \tag{6.53}$$

We have witnessed this phase relationship between quantum mechanical operators and Lie generators before in the case of the rotation group, where the $\hat{X}_k$ operators were the three rotation operators and the $\hat{Y}_k$ operators were the angular momentum operators (including the constant $\hbar$ to account for the unit of momentum):

$$\hat{L}_i = i\hbar\hat{X}_i. \tag{6.54}$$

The importance of this phase change follows from the fact that Hermitian operators correspond to *physical observables* in quantum mechanics. We will return to this point in § 6.3.2.

We end this section by reconsidering the invariance of observables, introduced in §6.1.3. Supposing that Eq. (6.12) holds true, we have, using Eq. (6.33):

$$\left[\hat{A}, \hat{T}\right] = \left[\hat{A}, \exp\left(\sum_{i=1}^{n} r_i \hat{X}_i\right)\right] = 0. \tag{6.55}$$

Since this must be the case for any r_i, we have

$$\left[\hat{A}, \hat{X}_i\right] = 0, \qquad \forall i = 1, \ldots, n. \tag{6.56}$$

Specializing to the Hamiltonian case as in §6.1.4, we have

$$\left[\hat{\mathscr{H}}, \hat{X}_i\right] = 0, \qquad \forall i = 1, \ldots, n. \tag{6.57}$$

Definition 6.8 (Invariance of the Hamiltonian): Let $\hat{\mathscr{H}}$ denote the Hamiltonian of a quantum system $\mathscr{S}$, which is invariant under the action of a *symmetry operator* $\hat{T}(\boldsymbol{\rho})$, with $\boldsymbol{\rho} = \{r_1, r_2, \ldots, r_n\}$; then, $\hat{\mathscr{H}}$ *commutes* with the n generators of $\hat{T}$: $\left[\hat{\mathscr{H}}, \hat{X}_i\right] = 0$ $(i = 1, \ldots, n)$. ▪

6.3 SYMMETRY AND CONSERVATION LAWS

6.3.1 Noether's theorem

Summarizing the previous sections, we have seen that when the Hamiltonian $\hat{\mathscr{H}}$ of a quantum system is invariant under a unitary transformation $\hat{T}$, generated by the Hermitian operator $\hat{Y}$, then $\hat{\mathscr{H}}$ commutes with both $\hat{T}$ and $\hat{Y}$ (i.e., $\left[\hat{\mathscr{H}}, \hat{T}\right] = \left[\hat{\mathscr{H}}, \hat{Y}\right] = 0$). The physical observable Y, corresponding to the generator $\hat{Y}$, is then conserved in time—in other words, $d\langle\hat{Y}\rangle/dt = 0$, and Y is said to be a constant of the motion.

In short, *symmetries give rise to conserved quantities*. Conversely, each conserved quantity implies that the system is invariant under a group of continuous symmetry transformations. This is easily illustrated in the Hamiltonian context of quantum mechanics (see §§6.2.3–6.3.2). Yet, this revolutionary idea first originated in the Lagrangian context of classical mechanics through the work of mathematician Emmy Noether (1882–1935), who first derived it in the year 1915.[10] It is therefore known as *Noether's first theorem*:

Theorem 6.3 (Noether's first theorem): For every continuous symmetry property of a dynamical system, there corresponds a quantity with a value that is conserved in time, and vice versa. ▪

According to American physicists Leon M. Lederman and Christopher T. Hill, Noether's theorem is "certainly one of the most important mathematical theorems ever proved in guiding the development of modern physics, possibly on a par with the Pythagorean theorem."[11] Albert Einstein similarly referred to Fräulein Noether as "the most significant creative mathematical genius thus far produced since the higher education of women began."[12]

The implications of Noether's theorem for our understanding of the laws of Nature are far-reaching. Physical experiments are generally assumed to have the same outcome regardless of their *location* and *orientation in space*, or their *position in time*. Whether an apple falls from a tree in Cambridge, England, or in Central Park in New York, Newton's laws of motion describe the downward fall in exactly the same

---

[10] E. Noether. "Invariante Variationsprobleme." *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen Mathematisch-Physikalische Klasse* (1918), pp. 235–257.

[11] L. M. Lederman and C. T. Hill. *Symmetry and the Beautiful Universe*. Amherst: Prometheus Books, 2004, p. 73.

[12] A. Einstein. "Professor Einstein Writes in Appreciation of a Fellow-Mathematician." *New York Times* May 5, 1935.

way. In a similar vein, it does not matter whether the apple fell back in the late 1660s or just an hour ago.

This implies that the laws of physics must be *invariant* (*symmetric*) under spatial translations ($\mathbf{a}' = \mathbf{a} + \mathbf{c}$), time translations ($t' = t + c$), and rotations ($\mathbf{a}' = \mathbb{R}(\omega \boldsymbol{n})\,\mathbf{a}$). Without these invariances, the scientific method would crumble in the light of the *irreproducibility* of experiments. Space translation symmetry is a consequence of the *homogeneity of space*, time translation symmetry relates to the *homogeneity of time*, and rotation symmetry is a result of the *isotropy of space*.

Now, following Noether, there is a *conservation law* (or *conserved quantity*) for each of these *continuous space–time symmetries*. Space translation symmetry gives rise to the conservation of linear momentum $\mathbf{p}$, time translation symmetry accounts for the conservation of energy E, and rotation symmetry allows for the conservation of angular momentum $\mathbf{L}$, as we saw in the previous chapter.

Notice that we are focusing here on the symmetries of *physical laws* and not on the symmetries of the physical *objects* or systems subject to these laws. In our example of the falling apple, for example, the apple itself does not need to be symmetric; rather, the physical laws governing the downward fall of the apple have to be symmetric. In this case, these are Newton's laws of motion. Because we are dealing primarily with quantum mechanics, the law that interests us most is the *time-dependent Schrödinger equation* $i\hbar\partial\psi/\partial t = \hat{\mathscr{H}}\psi$, which describes the evolution of a quantum system, just as Newton's laws of motion describe the evolution of a classical system, such as Newton's apple. This implies we will be mostly concerned with the symmetries of the Hamiltonian expression $\hat{\mathscr{H}}$.

6.3.2 Conserved quantities in quantum mechanics

There is a deep connection between the symmetry properties of a physical system and the *conservation laws* characterizing that system. To address this intimate relationship, it will be useful to express the *time-dependent Schrödinger equation* in both bra and ket space. In ket space,

$$i\hbar\frac{\partial\left|\psi\right\rangle}{\partial t} = \hat{\mathscr{H}}\left|\psi\right\rangle, \tag{6.58}$$

where $|\psi\rangle$ is the wave function of the system and $\hat{\mathscr{H}}$ is the Hamiltonian operator, which we assume to be independent of time. When written in the dual bra space (taking the Hermitian character of $\hat{\mathscr{H}}$ into account),

$$-i\hbar\frac{\partial\left\langle\psi\right|}{\partial t} = \left\langle\psi\right|\hat{\mathscr{H}}. \tag{6.59}$$

Let the operator $\hat{Y}$ correspond to a physical observable Y. For Y to be conserved, its expectation value $\langle\hat{Y}\rangle$ must not change in time. That is,

$$\frac{\partial\left\langle\hat{Y}\right\rangle}{\partial t} = 0. \tag{6.60}$$

With the help of the time-dependent Schrödinger equation in Eqs. (6.58) and (6.59), this can be rewritten as

$$\frac{\partial\left\langle\hat{Y}\right\rangle}{\partial t} = \frac{\partial}{\partial t}\left\langle\psi\right|\hat{Y}\left|\psi\right\rangle$$

$$= \left(\frac{\partial \langle\psi|}{\partial t}\right)\hat{Y}|\psi\rangle + \langle\psi|\left(\frac{\partial \hat{Y}}{\partial t}\right)|\psi\rangle + \langle\psi|\hat{Y}\left(\frac{\partial |\psi\rangle}{\partial t}\right)$$

$$= \frac{i}{\hbar}\langle\psi|\hat{\mathscr{H}}\hat{Y}|\psi\rangle + \langle\psi|\left(\frac{\partial \hat{Y}}{\partial t}\right)|\psi\rangle - \frac{i}{\hbar}\langle\psi|\hat{Y}\hat{\mathscr{H}}|\psi\rangle$$

$$= \langle\psi|\left(\frac{\partial \hat{Y}}{\partial t}\right)|\psi\rangle + \frac{i}{\hbar}\langle\psi|\left[\hat{\mathscr{H}},\hat{Y}\right]|\psi\rangle = 0.$$

Assuming $\hat{Y}$ to be independent of time, this expression reduces to

$$\frac{\partial\left\langle\hat{Y}\right\rangle}{\partial t} = \frac{i}{\hbar}\langle\psi|\left[\hat{\mathscr{H}},\hat{Y}\right]|\psi\rangle = 0. \tag{6.61}$$

It follows that $\partial\left\langle\hat{Y}\right\rangle/\partial t$ is zero (i.e., Y is conserved) when $\hat{\mathscr{H}}$ and $\hat{Y}$ commute:

$$\left[\hat{\mathscr{H}},\hat{Y}\right] = 0. \tag{6.62}$$

This is known as *Ehrenfest's theorem*, after Austrian physicist and mathematician Paul Ehrenfest (1880–1933).

Theorem 6.4 (Ehrenfest's theorem): Let $\hat{Y}$ represent a Hermitian operator. If $\hat{Y}$ commutes with the Hamiltonian $\hat{\mathscr{H}}$ (i.e., $\left[\hat{\mathscr{H}},\hat{Y}\right] = 0$), then Y is a conserved physical quantity, according to

$$\frac{\partial\left\langle\hat{Y}\right\rangle}{\partial t} = \frac{i}{\hbar}\langle\psi|\left[\hat{\mathscr{H}},\hat{Y}\right]|\psi\rangle = 0.$$

The physical observable Y is said to be a *constant of the motion.* ■

Since the Hamiltonian $\hat{\mathscr{H}}$ naturally commutes with itself (i.e., $\left[\hat{\mathscr{H}},\hat{\mathscr{H}}\right] = 0$), the energy E is conserved and is called a constant of the motion.[13] In view of Eq. (5.149), the components of the angular momentum L_i ($i = 1,2,3$) are constants of the motion as well. This implies that the total angular momentum **L** is invariant in time.

6.4 THE CARTAN-WEYL METHOD

Let $\mathfrak{g}$ represent a Lie algebra of dimension n spanned by the generators $\hat{X}_i$ ($i = 1,\ldots,n$) that satisfy the commutation rules

$$\left[\hat{X}_i,\hat{X}_j\right] = \sum_{k=1}^{n} f_{ijk}\hat{X}_k \equiv f_{ijk}\hat{X}_k, \tag{6.63}$$

where f_{ijk} are the structure constants. The n generators are said to form a *basis* $\mathcal{B}$ for the Lie algebra. It is possible to chose different bases for any particular Lie algebra $\mathfrak{g}$. In the case of the angular momentum algebra, for instance, a change of basis was made from $\left\{\hat{L}_x,\hat{L}_y,\hat{L}_z\right\}$ to $\left\{\hat{L}_+,\hat{L}_-,\hat{L}_z\right\}$.[14] One particularly useful basis is the *Cartan-Weyl basis*, which is introduced in this section.

---

[13] Although numerous constants of the motion will pass in review in the following chapters, *motion* itself does not always need to be involved, just evolution in time.

[14] Note that the structure constants are *basis dependent*. That is, a change of basis generally leads to different values for the structure constants.

The reason for devoting an entire section to this basis arises from the intimate relation between symmetries and degeneracies, as described in §6.1.6. That is, given the time-independent Schrödinger equation

$$\hat{\mathscr{H}}\,|\psi\rangle = E\,|\psi\rangle\,, \tag{6.64}$$

and invoking the commutation relationships $\left[\hat{\mathscr{H}}, \hat{X}_i\right] = 0$, $\hat{X}_i\,|\psi\rangle$ is an eigenket of $\hat{\mathscr{H}}$ as well:

$$\begin{aligned} \hat{X}_i\hat{\mathscr{H}}\,|\psi\rangle &= \hat{X}_i E\,|\psi\rangle \\ \hat{\mathscr{H}}\hat{X}_i\,|\psi\rangle &= E\hat{X}_i\,|\psi\rangle\,. \end{aligned} \tag{6.65}$$

The symmetry generators $\hat{X}_i$ can thus be used to relate the different eigenkets of a degenerate energy level. To this end, we traditionally replace the generators $\hat{X}_i$ by a linearly independent set of *Cartan* and *Weyl operators* $\hat{H}_i$ and $\hat{E}_\alpha$, which form the *Cartan-Weyl basis* $\{\hat{H}_i; \hat{E}_\alpha\}$ of $\mathfrak{g}$. Although the Cartan generators are used to *label* the different degenerate eigenstates, the Weyl generators act as *shift operators* or *ladder operators* between the different eigenkets, and enable us to *transform* one eigenket into another.

The construction of the Cartan-Weyl basis proceeds along a threefold path, as will be outlined in §6.4.1. Along the way, the concept of a *Weyl diagram* will also be introduced. Weyl diagrams provide a diagrammatic summary of the most important characteristics of any Lie algebra. And because the adage goes “A picture is worth a thousand words,” we can only urge you to read (and reread) this section with the greatest care and attention.

The rather formal treatment in §6.4.1 will be made more concrete in §6.4.2, where we will review some of the results from our study of the angular momentum algebra in light of the Cartan-Weyl method. This will form an ideal steppingstone for the treatment of more complex Lie algebras in the chapters that follow.

But, before we embark on our journey, let us briefly compare the defining relationships for a ladder operator (or eigenoperator) and an eigenfunction:

$$\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A} = c\hat{B}; \tag{6.66}$$

$$\hat{A}\,|a_i\rangle = a_i\,|a_i\rangle\,. \tag{6.67}$$

Just as $|a_i\rangle$ is called an *eigenfunction* of $\hat{A}$ with *eigenvalue* a_i in view of Eq. (6.67), $\hat{B}$ is said to be an *eigenoperator* of $\hat{A}$ with *eigenvalue* c in view of Eq. (6.66).

6.4.1 The threefold path toward a Cartan-Weyl basis

Step 1: The Cartan subalgebra and Cartan generators

In the first step, we find the *maximal subset of commuting generators* $\hat{X}_i$ $(i = 1, \ldots, m)$, which we will denote by the new symbol $\hat{H}_i$ for further convenience. The operators $\hat{H}_i$ form a basis for a *maximal Abelian subalgebra* $\mathfrak{H}$ of $\mathfrak{g}$, with a dimension $m < n$ that equals the number of commuting generators. Note that the subalgebra $\mathfrak{H}$ is maximal when there does not exist an Abelian subalgebra of higher dimension $\mathfrak{F} \subset \mathfrak{g}$ such that $\mathfrak{H} \subset \mathfrak{F}$.

Definition 6.9 (Maximal Abelian subalgebra): An Abelian subalgebra $\mathfrak{H}$ of a Lie algebra $\mathfrak{g}$ is said to be *maximal* when there are no additional elements of the algebra $\mathfrak{g}$ that commute with all the elements of the subalgebra $\mathfrak{H}$. ▪

The subalgebra $\mathfrak{H}$ is better known as the *Cartan subalgebra* of $\mathfrak{g}$, and the number of elements m in the Cartan subalgebra is called the *rank* of the Lie algebra $\mathfrak{g}$.

Definition 6.10 (Cartan subalgebra of a Lie algebra): Let $\mathfrak{g}$ denote an n-dimensional Lie algebra. Then, the set of all mutually commuting basis elements $\left\{\hat{X}_i = \hat{H}_i\right\}$ $(i = 1, \ldots, m)$ of $\mathfrak{g}$ forms a basis for a maximal Abelian subalgebra $\mathfrak{H}$ of $\mathfrak{g}$. This is called the *Cartan subalgebra* of $\mathfrak{g}$. ▪

Definition 6.11 (Rank of a Lie algebra): The dimension m of the Cartan subalgebra $\mathfrak{H} \subset \mathfrak{g}$ defines the *rank* of the Lie algebra $\mathfrak{g}$. ▪

The elements $\hat{H}_i$ of the Cartan subalgebra are called the *Cartan generators* or *Cartan elements.*[15] They satisfy the simple commutation relations:

$$\left[\hat{H}_i, \hat{H}_j\right] = 0, \qquad \forall i, j = 1, \ldots, m. \tag{6.68}$$

This means all $\hat{H}_i$ are simultaneously diagonalizable. Denoting their eigenvalues by h_i, we obtain the following eigenvalue equations:

$$\hat{H}_i \left| h_1, h_2, \ldots, h_m \right\rangle = h_i \left| h_1, h_2, \ldots, h_m \right\rangle, \qquad \forall i = 1, \ldots, m. \tag{6.69}$$

The eigenvalues h_i are called *weights* and can be considered components of an m-dimensional vector $\boldsymbol{h}$, which is termed the *weight vector.* The weights of a Cartan subalgebra can thus be used as quantum numbers to label the substates of a given multiplet.

Step 2: Weyl generators and Weyl diagrams

In the second step, we combine the remaining generators $\hat{X}_i$ of $\mathfrak{g}$ $(i = 1, \ldots, n - m)$ that are *not* elements of the Cartan subalgebra $\mathfrak{H}$ into linear combinations to form a linearly independent set of *raising* and *lowering* operators. We will denote these *ladder operators* by the general symbol $\hat{E}_\alpha$. The different $\hat{E}_\alpha$'s are also called *Weyl generators* or *Weyl elements.*

Along with the Cartan generators $\hat{H}_i$, they constitute the *Cartan-Weyl basis* for the Lie algebra $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ has actually been decomposed into a direct sum of the Cartan subalgebra $\mathfrak{H}$ (spanned by the m $\hat{H}_i$'s) and $n - m$ one-dimensional subalgebras $\mathfrak{E}_\alpha$, generated by the Weyl generators $\hat{E}_\alpha$:

$$\mathfrak{g} = \mathfrak{H} \bigoplus_{\alpha=1}^{n-m} \mathfrak{E}_\alpha = \mathfrak{H} \oplus \mathfrak{E}_1 \oplus \mathfrak{E}_2 \oplus \ldots \oplus \mathfrak{E}_{n-m}. \tag{6.70}$$

In view of their function as step operators, the $\hat{E}_\alpha$'s are *eigenoperators* of the Cartan generators $\hat{H}_i$. That is, they satisfy the following commutation relation:

$$\left[\hat{H}_i, \hat{E}_\alpha\right] = \alpha_i \hat{E}_\alpha, \qquad \forall i = 1, \ldots, m, \ \alpha = 1, \ldots, \ n - m. \tag{6.71}$$

---

[15] The $\hat{H}_i$'s are also called *Abelian generators* or *diagonal operators.*

We can rewrite this as

$$\begin{bmatrix} \left[\hat{H}_1, \hat{E}_\alpha\right] \\ \left[\hat{H}_2, \hat{E}_\alpha\right] \\ \vdots \\ \left[\hat{H}_m, \hat{E}_\alpha\right] \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \hat{E}_\alpha = \boldsymbol{\alpha} \hat{E}_\alpha. \tag{6.72}$$

The different α_i's are called the *roots* of $\hat{E}_\alpha$; the set of α_i's can be considered the components of a vector $\boldsymbol{\alpha}$, called the *root vector*, which lies in an m-dimensional *root space*.[16]

We can thus display the root vector for every Weyl generator $\hat{E}_\alpha$ in a diagram, which leads to a particularly powerful geometric representation of the Lie algebra $\mathfrak{g}$. This is called a *root diagram* or a *Weyl diagram*, the dimension of which is defined by the rank m of the Lie algebra. Notice that all Cartan generators $\hat{H}_i$ have their roots $\alpha_i = 0$ in view of Eq. (6.68) and lie at the center of the graph.

Naturally, root diagrams for rank $m \geq 3$ are difficult to visualize. For this reason, an alternative method of displaying root diagrams was introduced by Russian mathematician Eugene Borisovich Dynkin (1924–2014). However, because the rank of the Lie algebras considered in this book will never exceed the number 3, it will be unnecessary for us to delve into the symbolism of *Dynkin diagrams*.

Step 3: Casimir operators

In the third and final step, we can construct a number of *Casimir operators* $\hat{C}_\mu$ that commute with all the operators $\hat{X}_i$:

$$\left[\hat{C}_\mu, \hat{X}_i\right] = 0, \qquad \forall \mu = 1, \ldots, m,\ i = 1, \ldots, n. \tag{6.73}$$

The number of Casimir operators for a given Lie algebra $\mathfrak{g}$ is given by the rank m of that Lie algebra. This is also called *Racah's theorem*,[17] which is stated here without further proof.

Theorem 6.5 (Racah's theorem): For every Lie algebra $\mathfrak{g}$ of rank m, there are a total of m Casimir operators $\hat{C}_\mu$ ($\mu = 1, \ldots, m$) that commute with the generators $\hat{X}_i$ ($i = 1, \ldots, n$) of $\mathfrak{g}$.

As a consequence, all $\hat{C}_\mu$'s also commute with the Cartan generators $\hat{H}_i$:

$$\left[\hat{C}_\mu, \hat{H}_i\right] = 0, \qquad \forall \mu, i = 1, \ldots, m. \tag{6.74}$$

Therefore, we can find a complete set of states that are simultaneous eigenstates of all $\hat{C}_\mu$ and $\hat{H}_i$. Denoting the eigenvalues of $\hat{C}_\mu$ and $\hat{H}_i$ by c_μ and h_i, respectively, we can

---

[16] Notice the difference between the roots α_i, which are associated eigenvalues of the *eigengenerators* $\hat{E}_\alpha$, and the weights h_i, which are eigenvalues of *eigenfunctions*.

[17] G. Racah. "Sulla caratterizzazione delle rappresentazioni irriducibili dei gruppi semisemplici di Lie." *Atti della Accademia Nazionale dei Lincei, Rendiconti Classe dei Scienze fisiche, matematiche e naturali* 8 (1950), pp. 108–112.

represent the corresponding eigenstates as $|c_1, c_2, \ldots, c_m; h_1, h_2, \ldots, h_m\rangle$, or $\left|c_\mu; h_i\right\rangle$ for short. This gives rise to the following eigenvalue equations:

$$\hat{C}_\mu \left|c_\mu; h_i\right\rangle = c_\mu \left|c_\mu; h_i\right\rangle, \tag{6.75}$$

$$\hat{H}_i \left|c_\mu; h_i\right\rangle = h_i \left|c_\mu; h_i\right\rangle, \tag{6.76}$$

for all $\mu, i = 1, \ldots,\ m$. In conclusion, the Casimir operators $\hat{C}_\mu$ and the generators $\hat{H}_i$ of the Cartan subalgebra allow us to label every state of a multiplet. The Ladder operators $\hat{E}_\alpha$, on the other hand, enable us to move between the states within that multiplet, as illustrated in a Weyl diagram. That is, when a Weyl generator $\hat{E}_\alpha$ acts on the ket $\left|c_\mu; h_i\right\rangle$, it shifts the eigenvalue of the operators $\hat{H}_i$ by an amount α_i according to

$$\hat{E}_\alpha \left|c_\mu; h_i\right\rangle \sim \left|c_\mu; h_i + \alpha_i\right\rangle, \tag{6.77}$$

analogous to the raising and lowering operators $\hat{L}_+$ and $\hat{L}_-$ of the angular momentum algebra. This can be seen as follows, using Eqs. (6.71) and (6.76):

$$\begin{aligned} \hat{H}_i \hat{E}_\alpha \left|c_\mu; h_i\right\rangle &= \left(\left[\hat{H}_i, \hat{E}_\alpha\right] + \hat{E}_\alpha \hat{H}_i\right) \left|c_\mu; h_i\right\rangle \\ &= \left(\alpha_i \hat{E}_\alpha + \hat{E}_\alpha h_i\right) \left|c_\mu; h_i\right\rangle \\ &= \left(h_i + \alpha_i\right) \hat{E}_\alpha \left|c_\mu; h_i\right\rangle, \end{aligned} \tag{6.78}$$

which leads to Eq. (6.77).

6.4.2 Review of the angular momentum algebra

The Cartan subalgebras and Weyl diagrams just introduced will reveal themselves as crucial tools in our analysis of the Lie groups and corresponding Lie algebras of the hydrogen atom and the periodic system. The importance of these concepts cannot be stressed enough, and we therefore end this chapter by reviewing some of the more important properties of the *angular momentum algebra*, illustrating the principles of §6.4.1 along the way.

The well-known commutation relations between the components of the angular momentum operator $\hat{L}$ are given by

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar \hat{L}_z, \qquad \left[\hat{L}_y, \hat{L}_z\right] = i\hbar \hat{L}_x, \quad \text{and} \quad \left[\hat{L}_z, \hat{L}_x\right] = i\hbar \hat{L}_y. \tag{6.79}$$

These commutation rules are said to define the *angular momentum algebra* (see §5.6.1), which is isomorphic to the $\mathfrak{so}(3)$ Lie algebra. Notice that none of the $\hat{L}_i$ operators commutes among each other. To disclose the consequences of this algebra, we select the $\hat{L}_z$ operator as the only *Cartan generator*, forming the *Cartan subalgebra* $\mathfrak{H}$. The square of the total angular momentum,

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \tag{6.80}$$

commutes with all the angular momentum operators

$$\left[\hat{L}^2, \hat{L}_x\right] = \left[\hat{L}^2, \hat{L}_y\right] = \left[\hat{L}^2, \hat{L}_z\right] = 0, \tag{6.81}$$

and, by virtue of this property, $\hat{L}^2$ is called a *Casimir operator* of the angular momentum algebra (see §5.6.2). Because $\mathfrak{H}$ is one-dimensional (i.e., the $\mathfrak{so}(3)$ algebra is of *rank* 1), this is the only Casimir operator of the $\mathfrak{so}(3)$ algebra (Theorem 6.5).

The commutation relations in Eqs. (6.79) and (6.81) imply the existence of a complete set of states that are simultaneous eigenstates of $\hat{L}^2$ and $\hat{L}_z$ (see §5.6.3).

We denoted the *eigenvalues* of $\hat{L}^2$ and $\hat{L}_z$ by l and m_l, respectively, and represent the corresponding *eigenstates* as $|l, m_l\rangle$. This gives rise to the following eigenvalue equations:

$$\hat{L}^2 \, |l, m_l\rangle = l\,(l+1)\,\hbar^2 \, |l, m_l\rangle; \tag{6.82}$$

$$\hat{L}_z \, |l, m_l\rangle = m_l \hbar \, |l, m_l\rangle \,. \tag{6.83}$$

Ladder operators

The remaining two operators, $\hat{L}_x$ and $\hat{L}_y$, are then combined into the following linear combinations (see §5.6.4):

$$\hat{L}_\pm \equiv \hat{L}_x \pm i\hat{L}_y. \tag{6.84}$$

The introduction of $\hat{L}_+$ and $\hat{L}_-$ gives rise to a number of new commutation relations. First of all, because $\hat{L}^2$ commutes with all the angular momentum operators, it also commutes with $\hat{L}_+$ and $\hat{L}_-$:

$$\left[\hat{L}^2, \hat{L}_+\right] = \left[\hat{L}^2, \hat{L}_-\right] = 0. \tag{6.85}$$

Of greater importance are the commutation relations of $\hat{L}_+$ and $\hat{L}_-$ with the Cartan generator $\hat{L}_z$:

$$\left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hbar\hat{L}_\pm. \tag{6.86}$$

Since Eq. (6.86) satisfies the defining operator relation in Eq. (6.71) for *ladder operators* (see also Theorem 5.2), $\hat{L}_+$ and $\hat{L}_-$ are called *Weyl generators*. Along with the Cartan generator $\hat{L}_z$, they constitute the *Cartan-Weyl basis* of the $\mathfrak{so}(3)$ algebra.

In view of Eq. (6.86), the operation of $\hat{L}_+$ or $\hat{L}_-$ on one of the simultaneous eigenkets of $\hat{L}^2$ and $\hat{L}_z$ transforms this eigenket $|l, m_l\rangle$ into another eigenket of $\hat{L}^2$ and $\hat{L}_z$ with the same eigenvalue l for $\hat{L}^2$, but with the eigenvalue m_l raised or lowered by one:

$$\hat{L}_\pm \, |l, m_l\rangle \rightarrow |l, m_l \pm 1\rangle \,. \tag{6.87}$$

By virtue of these properties, $\hat{L}_+$ is called a *raising operator* and $\hat{L}_-$ is called a *lowering operator*. Starting from a given eigenket $|l, m_l\rangle$, a ladder of eigenkets can thus be produced by the repeated application of $\hat{L}_+$ and $\hat{L}_-$:

$$\ldots \overset{\hat{L}_\pm}{\longleftrightarrow} |l, m_l - 1\rangle \overset{\hat{L}_\pm}{\longleftrightarrow} |l, m_l\rangle \overset{\hat{L}_\pm}{\longleftrightarrow} |l, m_l + 1\rangle \overset{\hat{L}_\pm}{\longleftrightarrow} \ldots. \tag{6.88}$$

However, it turns out that the values of m_l are bounded from above and below by the value of l, ranging over the interval $\{-l,\ -l+1,\ \ldots,\ l-1,\ l\}$. To every eigenvalue l, there thus corresponds a *multiplet* of $2l+1$ eigenkets, which are interconnected via the ladder operators $\hat{L}_+$ and $\hat{L}_-$:

$$|l, -l\rangle \overset{\hat{L}_\pm}{\longleftrightarrow} |l, -l+1\rangle \overset{\hat{L}_\pm}{\longleftrightarrow} \ldots \overset{\hat{L}_\pm}{\longleftrightarrow} |l, l-1\rangle \overset{\hat{L}_\pm}{\longleftrightarrow} |l, l\rangle \,. \tag{6.89}$$

This can also be represented *diagramatically*. To this aim, let us use the m_l values as coordinates to plot the $2l+1$ substates of a given multiplet on the axis of $\hat{L}_z$ eigenvalues (in units of $\hbar$). In such *weight diagrams*, a given multiplet thus appears as a line with a set of equidistant points, symbolizing the ladder of eigenstates. An example is shown in Figure 6.4B for the multiplet with $l = 5/2$.

By choosing the z-axis as the axis of quantization, we have placed the $\hat{L}_z$ operator in the center of the angular momentum algebra. That is, the diagonal operator $\hat{L}_z$ has been used to identify all components of a given multiplet.

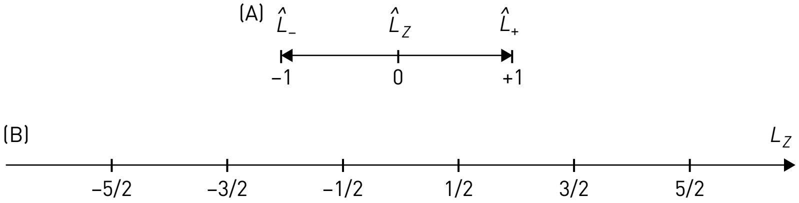


FIGURE 6.4 (A) Root diagram of the $\mathfrak{so}(3)$ algebra. The $\hat{L}_z$ operator lies in the center of the graph and is used to *label* the eigenkets of a given multiplet. The $\hat{L}_+$ and $\hat{L}_-$ operators act as shift operators. (B) Weight diagram for the multiplet with $l = 5/2$. Each eigenket is depicted by a point on the $\hat{L}_z$-axis that denotes its m_l-value.

Eigenoperators

We can go one step further and represent the generators themselves in a diagram. For this, we need to recall the concept of an *eigenoperator*. It is clear that the Cartan generator $\hat{L}_z$ is an eigenoperator of itself, with eigenvalue 0:

$$\left[\hat{L}_z, \hat{L}_z\right] = 0\hat{L}_z. \tag{6.90}$$

The operators $\hat{L}_x$ and $\hat{L}_y$, on the other hand, are *not* eigenoperators of $\hat{L}_z$ since their commutators do not return $\hat{L}_x$ and $\hat{L}_y$, respectively:

$$\left[\hat{L}_z, \hat{L}_x\right] = +i\hbar\hat{L}_y; \tag{6.91}$$

$$\left[\hat{L}_z, \hat{L}_y\right] = -i\hbar\hat{L}_x. \tag{6.92}$$

To find proper eigenoperators, we need to make linear combinations of the x- and y-components. The resulting combinations are precisely the raising and lowering operators $\hat{L}_+$ and $\hat{L}_-$:

$$\left[\hat{L}_z, \hat{L}_+\right] = +\hbar\hat{L}_+; \tag{6.93}$$

$$\left[\hat{L}_z, \hat{L}_-\right] = -\hbar\hat{L}_-. \tag{6.94}$$

The corresponding eigenvalues for the $\hat{L}_+$ and $\hat{L}_-$ eigenoperators are, respectively, $+\hbar$ and $-\hbar$, as is evident from the previous equations. The generators of the angular momentum algebra are thus represented by three points (or *roots*) with coordinates -1, 0, and $+1$ (in units of $\hbar$) in a one-dimensional *root space*. The resulting *root diagram* with *root vectors* for $\hat{L}_+$ and $\hat{L}_-$ is shown in Figure 6.4A.

The $\hat{L}_z$ operator forms an $\mathfrak{so}(2)$ subalgebra on its own because it generates an infinitesimal rotation around the z-axis. Of course, in the current case, this Cartan subalgebra is rather trivial because it contains only one element. In subsequent chapters we will meet extended quantization schemes in which several quantum numbers are needed to characterize shell components.

Reducing direct products of multiplets

We end this section with an application of the Weyl diagrams, which will be of great use in the following chapters—namely, the combination of different multiplets. Consider a system, consisting of several parts, in which each part is characterized by a multiplet with a given l-value. An example is a system with both orbital and spin angular momentum. The multiplets of the parts can then be combined into the multiplet of the whole system. To this aim, the *direct product* (or *tensor product*) has to be formed between the constituent multiplets. This is particularly easy when use is made of Weyl diagrams.

Let us see how it is done. Suppose, for the sake of illustration, that we want to form the direct product between a p- and a d-multiplet with $l_1 = 1$ and $l_2 = 2$, respectively. The p-multiplet has $2l_1 + 1 = 3$ substates with $m_l = -1,\ 0,\ 1$; and the d-multiplet has $2l_2 + 1 = 5$ substates with $m_l = -2,\ -1,\ 0,\ 1,\ 2$.

The formation of the direct product can be represented in a graphical way by superimposing a p-multiplet on top of each member state of the d-multiplet, as shown in the upper part of Figure 6.5. In the end, as shown in the lower part, $3 \times 5 = 15$ states are formed, with the L_z-values ranging between -3 and $+3$. That is, in general, the m_l-values range over the interval $\{-(l_1 + l_2),\ \ldots,\ l_1 + l_2\}$.

The composite system can then be decomposed as the *direct sum* of three multiplets: a p-, d- and f-multiplet. We write this symbolically as

$$p \otimes d = p \oplus d \oplus f, \tag{6.95}$$

where $\otimes$ denotes direct product and $\oplus$ indicates direct sum. Or, in terms of their multiplet dimension:

$$[3] \otimes [5] = [3] \oplus [5] \oplus [7]. \tag{6.96}$$

6.5 THE THREE PILLARS OF GROUP THEORY

Let us attempt to summarize, in a couple of lines, what has been introduced in this chapter. In a nutshell, we have seen that a *constant of the motion X* is a telltale sign for the existence of a particular *symmetry group* $\mathcal{G}$ of the system $\mathscr{S}$, which in turn can be invoked to rationalize the *degeneracies* observed in the energy eigenspectrum of $\mathscr{S}$. We could reason the other way around as well, starting from an observed *degeneracy*

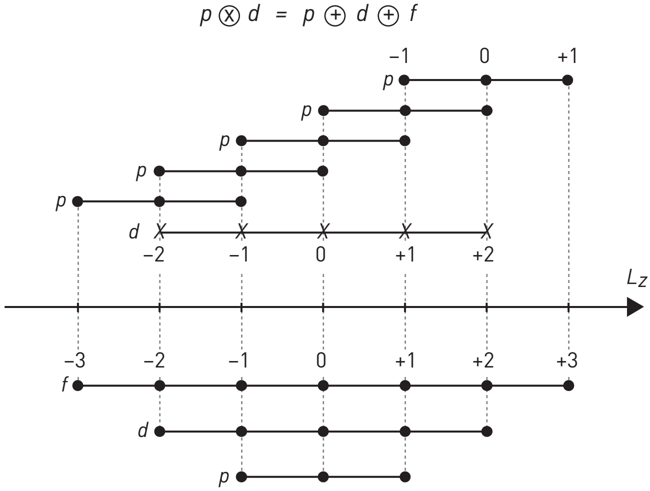


FIGURE 6.5 Weight diagrams illustrating the reduction of the tensor product $p \otimes d = p \oplus d \oplus f$, which yields the global multiplets for the composite system consisting of two parts.

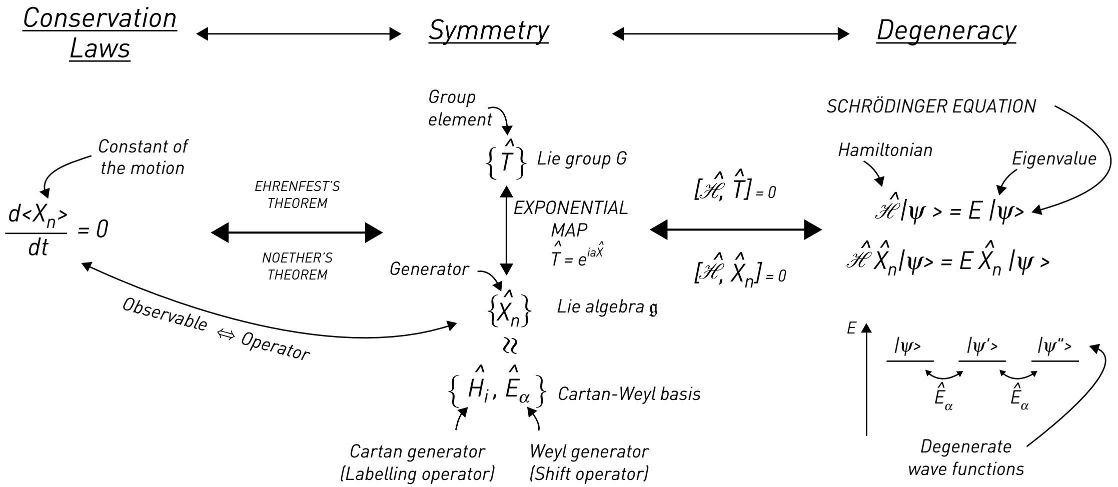


FIGURE 6.6 Conservation laws, symmetries, and degeneracy.

and looking for its corresponding *symmetries* and *conservation laws*. In short, we are faced with the following important interconnections:

$$\text{Conservation Law} \longleftrightarrow \text{Symmetry} \longleftrightarrow \text{Degeneracy}. \qquad (6.97)$$

As should be clear from this and the schematic depiction in Figure 6.6, the conservation laws, symmetries, and degeneracies of a system are not to be seen as independent or disconnected. Rather, they form a closely woven tapestry, with the theorems of Noether, Ehrenfest, and Wigner acting as a metaphorical Ariadne's thread interweaving these three pillars into a harmonious and powerful whole. Not surprisingly, then, the relations in Eq. (6.97) will permeate the presentation of the elementary particle zoo and the hydrogen atom in the next few chapters, and they will figure prominently in our discussion of the hidden symmetries of the periodic system in Chapter 13.

PART TWO

DYNAMIC SYMMETRIES

7 The SU(3) group

The world of the quark has everything to do with a jaguar circling in the night.
–Arthur Sze (1987)[1]

7.1 HISTORICAL PRELUDE: THE PARTICLE ZOO

7.1.1 Probing the inner structure of the atom (1897–1947)

Since the musings of the ancient Greek philosophers in the sixth century BC, scientists have pondered incessantly over the question: What are the most basic constituents of matter? Around 1947, physicists believed they had finally reached a simple, yet elegant, answer to this fundamental question.[2] They proposed that all matter around us was composed of *atoms*, which acted as miniature solar systems. The positive charge and most of the mass was thought to be concentrated in a tiny central *nucleus* (the sun), with the electrons (planets) swirling around the nucleus and held in orbit via

---

[1] From Arthur Sze's poem "The Leaves of a Dream Are the Leaves of an Onion," which first appeared in his collection *River, River* in 1987. Quoted in M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex*. New York: W. H. Freeman, 1994, p. 11.

[2] For a brief history of elementary particle physics, see D. Griffiths. *Introduction to Elementary Particles*. Weinheim: Wiley-VCH, 2008, pp. 11–53. A. Pais. *Inward Bound: Of Matter and Forces in the Physical World*. Oxford: Oxford University Press, 1988; M. Veltman. *Facts and Mysteries in Elementary Particle Physics*. Singapore: World Scientific, 2003. See also N. Mee. *The Higgs Force, the Symmetry-Breaking Force That Makes the World an Interesting Place*. Cambridge: Lutterworth Press, 2012, pp. 167–204.

the electromagnetic forces between the positively charged nucleus and the negatively charged electrons.

Hydrogen ($^1_1$H), for instance, was considered to be a two-body system with one *electron* (e^-) encircling the $+1$ positively charged nucleus, which Ernest Rutherford termed the *proton* (denoted p^+). With James Chadwick's discovery of the *neutron* (n^0) in 1932, the second constituent of the atomic nucleus was finally identified. Note that the mass[3] of the neutron (939.565378 MeV/c^2), is comparable to the proton mass (938.272046 MeV/c^2), but both are substantially heavier than the electron, which has a feathery mass of only 511.007 keV/c^2.

Three other elementary particles were known in the classical era. There was Einstein's *photon*, γ, and Wolfgang Pauli's *neutrino*, ν—a light-weight, electrically neutral particle, the existence of which was predicted by Pauli to account for the apparent loss of energy and momentum in radioactive beta decay. The third particle originated in Dirac's attempt to unify quantum theory with Einstein's theory of special relativity. Solutions of the Dirac equation hinted at the existence of *antiparticles*; for every particle there is an associated antiparticle with exactly the same mass but opposite electric charge. With the discovery of the *positron* e^+ (or antielectron) by Carl D. Anderson in 1932, Dirac's theory gained a first experimental verification.[4] Antiparticles are usually indicated by writing a bar above the symbol of their corresponding particle. An antiproton, for instance, is denoted by the symbol $\bar{p}$; an antineutron, by the symbol $\bar{n}$.[5]

In 1936, Anderson and his colleague Seth Neddermeyer also discovered a particle with a mass in between that of the electron and the proton, which they called a *mesotron*. Only much later was it recognized that this particle was not associated with the strong forces inside the nucleus, but that it belonged to the same family as the electron. It was renamed the *muon*.

7.1.2 Yukawa's pion (1947–1949)

The classical era came to an abrupt end in 1947 with the discovery of the *pion* (shorthand for *pi-meson*, denoted π; see also Table 7.1). The first pions were discovered in cosmic ray experiments performed by Cecil Powell and coworkers on the Pic du Midi de Bigorre in the Pyrenees. Photographic emulsions were placed for several hours on the mountain top before being developed. The impact of high-energy subatomic particles from the atmospheric cosmic rays left tracks on the photographic plates that could then be studied with a microscope to determine the mass and charge of the detected particles. With the first hints of the pion, Powell felt as if he had "broken

---

[3] Given Einstein's famous equation $E = mc^2$, the masses m of elementary particles are usually given in terms of their rest energy E (in keV, MeV, GeV, or even TeV).

[4] Negatively charged *antiprotons* were first discovered in 1955 by Emilio Segrè and Owen Chamberlain; the *antineutron* was detected one year later, in 1956, by Bruce Cork at Berkeley.

[5] You might wonder in which sense an antineutron is any different from a neutron, given that both particles are electrically neutral. For one thing, the neutron is composed of quarks whereas the antineutron is made out of antiquarks (see §7.6). As we will see later, many elementary properties, besides electric charge, are inverted upon turning a particle into its corresponding antiparticle. Some neutral particles, such as photons or the eta mesons, are their own antiparticles.

Table 7.1 Table of the properties of the low-mass *pseudoscalar mesons* with spin 0 (the spin concept is described in Chapter 8). As a result of their integer spin, mesons are classified as *bosons*. All mesons listed in this table are unstable and have a lifetime of 10^{-8} seconds or less. As a general rule, the lower the mass of the particle, the longer lived it will be. This table contains a number of particle–antiparticle pairs: π^+– π^-, K^+– K^-, and K^0– $\bar{K}^0$. The particles π^0, η^0, and η'^0 are their own antiparticles. The pseudoscalar mesons have *baryon number* $B = 0$, and *hypercharge* Y equal to strangeness S. The rest masses have been rounded.

Name	*Symbol*	*Charge*	*Isospin*	*Isospin Component*	*Strangeness*	*Rest Mass*
		(Q)	(T)	(T_3)	(S)	(MeV/c^2)
Pions	π^+	+1	1	+1	0	140
	π^0	0	1	0	0	135
	π^-	−1	1	−1	0	140
Kaons	K^+	+1	1/2	+1/2	+1	494
	K^0	0	1/2	−1/2	+1	498
	K^-	−1	1/2	−1/2	−1	494
	$\bar{K}^0$	0	1/2	+1/2	−1	498
Eta meson	η^0	0	0	0	0	548
Eta prime meson	η'^0	0	0	0	0	958

into a walled orchard, where protected trees flourished and all kinds of exotic fruits ripened in great profusion."[6]

It turns out that pions have a mass of 140 MeV. This is 280 times the electron mass and approximately one seventh times the proton mass. This intermediate mass led to the pions being classified as the genuine *mesons* (or "middleweight" particles, see Tables 7.1 and 7.2, with the former mesotron being dubbed *meson*, in accordance with the Greek μεσος); electrons and neutrinos belonged to the *lepton* family (for "lightweight"), and protons and neutrons were classed among the *baryons* (or "heavyweight"). Mesons and baryons together constituted the *hadron* family (from the Greek ἁδρος for "large" or "massive").

Pions come in three varieties: the positively charged π^+, the negatively charged π^-, and the neutral π^0. According to Hideki Yukawa's theory, all these pions were the *mediators* of the strong nuclear force. They were responsible for holding the protons and neutrons together in the nucleus, and counteracted the electromagnetic repulsion of the protons resulting from their positive electric charge. (In a similar vein, photons are said to mediate the electromagnetic force between the protons and electrons in an atom.) Interestingly, pions were observed to disintegrate into yet another particle, the *muon* μ; for example:

$$\begin{aligned} \pi^+ &\to \mu^+ + \nu; \\ \pi^- &\to \mu^- + \nu. \end{aligned} \tag{7.1}$$

---

[6] Quoted from C. F. Powell. "Fragments of Autobiography." In: *Selected Papers of Cecil Frank Powell.* Ed. E. H. S. Burhop, W. O. Lock, and M. G. K. Menon. New York: North-Holland, 1972, p. 26.

Table 7.2 Table of the properties of the *vector mesons* of spin 1. Because of their integer spin, vector mesons are also classified as *bosons*. (Numerous other mesons have been detected with even higher spins and masses.) As a result of the higher rest mass of most vector mesons as compared to the pseudoscalar mesons, vector mesons are generally highly unstable and have a lifetime of 10^{-20} s or less. This table contains a number of particle–antiparticle pairs: ρ^+– ρ^-, K^{*+}– K^{*-}, and K^{*0}– $\bar{K}^{*0}$. The particles ρ^0, ω^0, and ϕ^0 are their own antiparticles. The vector mesons have *baryon number* $B = 0$, and *hypercharge* Y equal to strangeness S. The rest masses have been rounded.

Name	*Symbol*	*Charge* (Q)	*Isospin* (T)	*Isospin Component* (T_3)	*Strangeness* (S)	*Rest Mass* (MeV/c^2)
Rho	ρ^+	+1	1	+1	0	775
mesons	ρ^0	0	1	0	0	775
	ρ^-	−1	1	−1	0	775
Kaon	K^{*+}	+1	1/2	+1/2	+1	892
resonances	K^{*0}	0	1/2	−1/2	+1	896
	K^{*-}	−1	1/2	−1/2	−1	892
	$\bar{K}^{*0}$	0	1/2	+1/2	−1	896
Omega	ω^0	0	0	0	0	783
Phi meson	ϕ^0	0	0	0	0	1019

The role of this lightweight particle, classified in the lepton family, was unknown at the time and led to Isidor Isaac Rabi's exclamation: "Who ordered *that*?"

7.1.3 The growing particle jungle (1950–1960)

With the construction of the first particle accelerators during the 1950s, such as the Brookhaven Cosmotron, the discovery of elementary particles was no longer dependent on cosmic ray experiments. The first synchrotrons reached energies in the gigaelectronvolt range. After two protons had been accelerated to such velocities, they were forced in a head-on collision. During the impact, most of the kinetic energy was transformed into matter, according to Einstein's equation $E = mc^2$. Within the resulting particle debris, novel particles were disclosed by studying the tracks left on photographic emulsions or traced in bubble and cloud chambers.

The hunt for new particles soon bore fruit, and the French garden of elementary particles rapidly transformed into a dense and wild jungle without apparent order. Some of these newly discovered particles were the so-called K-mesons or *kaons*, which came in four varieties: the positively charged K^+ along with its antiparticle K^-, and the neutral K^0 with its antiparticle $\bar{K}^0$. Kaons typically decay in a number of pions; for instance,

$$\begin{aligned} K^0 &\rightarrow \pi^+ + \pi^-; \\ K^+ &\rightarrow \pi^+ + \pi^+ + \pi^-. \end{aligned} \tag{7.2}$$

As a consequence, kaons are roughly two to three times as heavy as the pion; they thus belong to the meson family of middleweight particles (see Table 7.1). Another particle,

Table 7.3 Table of the properties of the low-mass *baryons* of spin 1/2. As a result of their half-integer spin, baryons are classed among the *fermions*. The proton is believed to be a stable particle; the free neutron has a lifetime of approximately 15 minutes. Most of the other baryons listed in this table are shorter lived, with a lifetime of 10^{-10} s or less. The corresponding antiparticle symbols can be obtained by adding a bar over the particles' symbols, and by inverting the signs of all charges and strangeness quantum numbers. All baryons have *baryon number* $B = +1$; antibaryons have $B = -1$. The rest masses have been rounded.

Name	*Symbol*	*Charge*	*Isospin*	*Isospin Component*	*Strangeness*	*Hyper-charge*	*Rest Mass*
		(Q)	*(T)*	*(T_3)*	*(S)*	*(Y)*	*(MeV/c^2)*
Nucleons	p^+	+1	1/2	+1/2	0	1	938
	n^0	0	1/2	−1/2	0	1	940
Lambda	Λ^0	0	0	0	−1	0	1116
Sigmas	Σ^+	+1	1	+1	−1	0	1189
	Σ^0	0	1	0	−1	0	1193
	Σ^-	−1	1	−1	−1	0	1197
Xis	Ξ^0	0	1/2	+1/2	−2	−1	1315
	Ξ^-	−1	1/2	−1/2	−2	−1	1322

found by Anderson's group in 1950, was the *lambda* particle Λ^0, which transformed into a proton and a pion:

$$\Lambda^0 \rightarrow p^+ + \pi^-. \tag{7.3}$$

This decay mode provided an explanation for the substantially greater Λ^0 mass (1116 MeV/c^2) as compared to the mass of the proton (938 MeV/c^2), and thus convinced the scientific community to classify it among the baryons, along with the proton and the neutron (see Table 7.3).

During the following years, a plethora of other baryons were discovered: the *sigma* particles (Σ^+, Σ^0 and Σ^-, as well as Σ^{*+}, Σ^{*0} and Σ^{*-}), the cascade or *xi* particles (Ξ^0 and Ξ^-, as well as Ξ^{*0} and Ξ^{*-}), the *delta* particles (Δ^{++}, Δ^+, Δ^0, and Δ^-), and so on (see Tables 7.3 and 7.4).

7.1.4 Quantum numbers and conservation laws

Not only did the discovery of all these particles come as a big surprise, physicists were baffled by their peculiar behavior. During the course of the next few years, an overwhelming variety of different particle reactions were observed in the heart of particle accelerators. As a result, the central problem of particle physics became one of *understanding* why certain reactions took place whereas others did not. For example, although a proton and a pion had been observed to transform into a kaon (*meson*) and a lambda particle (*baryon*):

$$p^+ + \pi^- \rightarrow K^0 + \Lambda^0, \tag{7.4}$$

Table 7.4 Table of the properties of *baryon resonances*. These are extremely short-lived baryon states with lifetimes of 10^{-23} s or less. All baryon resonances listed in this table have spin $3/2$ and belong to the class of *fermions*. In complete analogy with the baryons of Table 7.3, the corresponding antiparticles can be obtained by adding a bar over the particles' symbols, and by inverting the signs of all charges and strangeness quantum numbers. All baryon resonances have *baryon number* $B = +1$; antibaryons have $B = -1$. The rest masses have been rounded.

Name	*Symbol*	*Charge*	*Isospin*	*Isospin Component*	*Strangeness*	*Hyper-charge*	*Rest Mass*
		(Q)	*(T)*	*(T_3)*	*(S)*	*(Y)*	*(MeV/c^2)*
Deltas	Δ^{++}	$+2$	$3/2$	$+3/2$	0	$+1$	1232
	Δ^{+}	$+1$	$3/2$	$+1/2$	0	$+1$	1232
	Δ^{0}	0	$3/2$	$-1/2$	0	$+1$	1232
	Δ^{-}	-1	$3/2$	$-3/2$	0	$+1$	1232
Sigmas	Σ^{*+}	$+1$	1	$+1$	-1	0	1385
	Σ^{*0}	0	1	0	-1	0	1385
	Σ^{*-}	-1	1	-1	-1	0	1385
Xis	Ξ^{*0}	0	$1/2$	$+1/2$	-2	-1	1530
	Ξ^{*-}	-1	$1/2$	$-1/2$	-2	-1	1530
Omega	Ω^{-}	-1	0	0	-3	-2	1672

the following two reactions never occurred:

$$p^+ + \pi^- \nrightarrow K^0 + \Sigma^-; \tag{7.5}$$

$$p^+ + \pi^- \nrightarrow K^0 + \Xi^0, \tag{7.6}$$

even though Σ^- and Ξ^0 are *baryons*, just like the Λ^0 particle. To tackle this problem, particle physicists adopted a *phenomenological approach* by imposing a series of *conservation laws* that forbade certain reactions from happening.

We can illustrate this approach by considering the *law of conservation of electric charge*, which says that electric charge can neither be created nor destroyed; that is, the net amount of charge in any system is a *conserved quantity*. Electric charge occurs in discrete quantities. In other words, charge is measured in units of the *elementary charge* e. A proton, for instance, has charge $+e$; an electron has charge $-e$. Dividing the charge of a particle by the elementary charge e yields the *charge number* Q. Thus, a proton has a charge number of $+1$ and an electron of -1 (cf. Table 7.3). The charge number Q is called a *quantum number*. The conservation of electric charge then amounts to the following statement:

$$Q_{\text{initial}} = Q_{\text{final}}, \tag{7.7}$$

with Q_{initial} and Q_{final} being the sum of Qs over all particles in the initial and final configurations, respectively. Reactions in which Q is not conserved (i.e., for which $Q_{\text{initial}} \neq Q_{\text{final}}$) are forbidden. In the first reaction (7.4), for instance, $Q_{\text{initial}} = (+1) + (-1) = 0$ and $Q_{\text{final}} = 0 + 0 = 0$. Because the electric charge is conserved, the reaction is allowed. The second reaction (7.5), on the other hand, is forbidden because the electric charge is *not* conserved in this case: $Q_{\text{initial}} = 0 \neq Q_{\text{final}} = -1$.

A strange phenomenon is seen to occur in reaction (7.6), however. Although the electric charge is conserved, the reaction is not allowed. As a result, physicists were forced to introduce new *quantum numbers*, such as baryon number, lepton number, electron number, muon number, and strangeness (see §7.1.5), to rationalize all observed (and nonobserved) particle reactions.

7.1.5 Strangeness (1953)

Another problem facing the theoretical physicists during the 1950s was connected to the decay of certain baryons. Although most baryons were *created* almost instantaneously (on a timescale of 10^{-23} s), some of them (such as the kaons or the lambda particle) were seen to *decay* considerably slower (on the order of 10^{-10} s). To explain this "strange" behavior, Murray Gell-Mann and Kazuhiko Nishijima independently introduced a new quantum number with the name *strangeness* (denoted with the symbol S). Just like electric charge is a conserved quantity in elementary particle interactions, Gell-Mann and Nishijima believed strangeness would be conserved in strong interactions; that is,

$$S_{\text{initial}} = S_{\text{final}}. \tag{7.8}$$

For example, in the proton–pion collision (7.4) of the previous paragraph, $p^+ + \pi^- \rightarrow K^0 + \Lambda^0$, both the proton and the pion have $S = 0$ (cf. Table 7.3), and because the kaon K^0 carries strangeness $S = +1$, Λ^0 must have $S = -1$. That is, for each particle with strangeness $+1$, formed under the strong force, another particle with strangeness -1 is observed to be created simultaneously. This also explains the *non*-occurrence of reaction (7.6): $p^+ + \pi^- \nrightarrow K^0 + \Xi^0$, where Ξ^0 has a strangeness $S = -2$ ($S_{\text{initial}} = 0 \neq S_{\text{final}} = -1$).

What is more, all particles with nonzero strangeness are highly unstable, and they would love to decay into particles with strangeness 0. This is not possible, however, because strangeness is conserved under the strong force. Gell-Mann and Nishijima therefore postulated that strangeness would no longer be conserved in *weak* interactions. In this way, strange particles (or *hyperons* as they were called) could decay under the weak force, albeit at a much slower rate. For example, in the decay of the lambda particle,

$$\Lambda^0 \rightarrow p^+ + \pi^-, \tag{7.9}$$

strangeness is clearly *not* conserved. The conservation of this fundamental property under the strong nuclear force hinted at a deep symmetry, as we intend to recount in this chapter.

7.1.6 The Mendeleev of elementary particle physics

By the beginning of the 1960s, physicists were gravely bothered by so many unexpected particles with startling properties. Elementary particle physics had turned into a new form of butterfly collecting. Deeply annoyed by this discomforting situation, Enrico Fermi told one of his students, "Young man, if I could remember the names of all these particles, I would have been a botanist." Willis Lamb in his Nobel Prize lecture said, "I have heard it said that the finder of a new elementary particle used to be rewarded by a Nobel Prize, but such a discovery now ought to be punished by a $10.000 fine."

The state of elementary particle physics at the beginning of the 1960s resembled the situation in chemistry a century before. During the 1860s, numerous chemical elements had been isolated and studied in the lab, but no apparent order was found among the different elements. It was the Russian chemist Dmitri Ivanovich Mendeleev (1834–1907) who finally came up with a periodic system in 1869 in which all the elements were classified along horizontal periods according to increasing atomic mass, and in vertical groups according to chemical similarities, thus exhibiting the various relationships among the different elements in a tabular format (see Chapter 13).

A century later, in 1961, American physicist Gell-Mann (b. 1929) stepped forward as the Mendeleev of elementary particle physics. Guided by a profound belief in the force of symmetry, Gell-Mann brought order in the apparent chaos of elementary particles. The clarification and systematization of particle physics brought about in his *eightfold way* (§7.5) culminated in his prediction of *quarks* (§7.6), which earned him a Nobel Prize in physics in 1969, exactly one hundred years after Mendeleev's discovery of the periodic system. For more than twenty years (from 1950–1970 onward), Gell-Mann dominated the field of theoretical particle physics and stamped his mark by introducing numerous catchy names such as "strangeness" (*vide supra*), the "eightfold way," and "quarks" (*vide infra*).[7]

There is a lot of ground to be covered before we can fully appreciate the power and elegance of Gell-Mann's insight. For one thing, we have to acquaint ourselves with a new form of symmetry group. The Lie group involved is the *special unitary group* in three dimensions, SU(3), which in many respects can be viewed as SO(3)'s older brother.

The best way of introducing the SU(3) group is via a consideration of the three-dimensional *isotropic harmonic oscillator* that shares the same symmetry as the eightfold way. The fact that the SU(3) group manifests itself in these two systems does not necessarily imply that there exists a deep relationship between both. It simply shows that "Nature seems to like the number three!" And yet, much of the machinery to be introduced in our treatment of the harmonic oscillator (§§7.2–7.3) proves applicable to the quark structure as well (§7.6). This is another example of how one and the same *abstract* group can be realized in two *concrete* and altogether different systems (see also Chapter 2, §2.2).

To achieve a *gradual* understanding of the SU(3) symmetry, let us start with the unitary group in one dimension, U(1), which is the symmetry of the one-dimensional harmonic oscillator (§7.2). From there, it is straightforward to generalize to the three-dimensional case with SU(3) symmetry (§7.3). Finally, after describing the general properties of the U(3) and SU(3) group in §7.4, the story of Gell-Mann's eightfold way will be recounted in §7.5. We will conclude this chapter with a brief journey to the subatomic realm of quarks in §7.6.

---

[7] For a biography of Murray Gell-Mann, see G. Johnson. *Strange Beauty: Murray Gell-Mann and the Revolution in Twentieth-Century Physics*. New York: Vintage Books, 2000. See also Gell-Mann's popular scientific book M. Gell-Mann. *The Quark and the Jaguar*. 1994. For a collection of papers relating to the eightfold way, see M. Gell-Mann and Y. Ne'eman. *The Eightfold Way*. Boulder, CO: Westview Press, 2000.

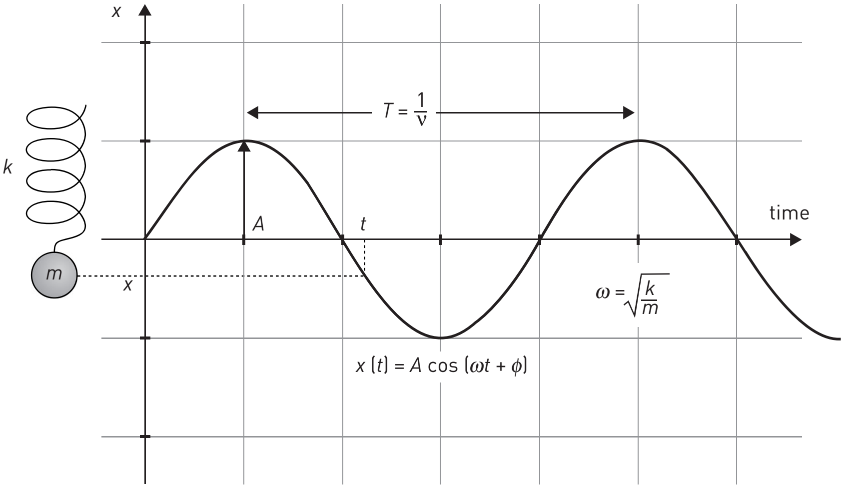


FIGURE 7.1 Simple harmonic motion of a mass m attached to a spring with spring constant k, yielding sinusoidal oscillations about the equilibrium $x = 0$, with A the amplitude, ν the frequency, T the period, and ω the angular velocity. Note that in this situation, the phase ϕ equals $-\pi/2$.

7.2 THE ONE-DIMENSIONAL HARMONIC OSCILLATOR

The *one-dimensional harmonic oscillator* is sometimes called "the model of all models." It is one of the simplest integrable systems with ubiquitous ramifications, from molecular vibrations and boson matter down to string theory. Our examination of this system starts during the eighteenth century—1747, to be precise—when Jean-Baptiste Le Rond d'Alembert, for the first time, described the wave equation of a vibrating string, which is essentially the classical mechanics of the one-dimensional harmonic oscillator. Another important pillar is Robert Hooke's equation, which dates from even earlier times, back in the seventeenth century.

7.2.1 Classical mechanical treatment

Suppose a single particle with mass m is attracted toward the origin by a force proportional to the particle's displacement x from the origin (Figure 7.1). According to Hooke's law,

$$F = -kx, \tag{7.10}$$

with the proportionality constant k denoting the so-called *force constant* and F being the force acting on the particle. Eq. (7.10) is obeyed by a mass attached to a spring, provided the spring is not stretched too much from its equilibrium position at $x = 0$. This system is called a *simple harmonic oscillator* because F is the only force acting on the particle.[8] The force constant k is also referred to as the *spring constant.*

---

[8] If a frictional force were present, the system would be described as a *damped harmonic oscillator*.

Equations of motion

Combining Newton's second law $F = ma$ with Hooke's law $F = -kx$ gives a second-order linear differential equation:

$$F = ma = m\frac{d^2x}{dt^2} = -kx, \tag{7.11}$$

where use was made of Eq. (7.10) and t represents time. This equation is essentially d'Alembert's wave equation. The solutions to this equation are given by

$$x(t) = A\cos(2\pi\nu t + \phi) = A\cos(\omega t + \phi), \tag{7.12}$$

and they describe simple harmonic motion of a periodic nature, (i.e., consisting of sinusoidal oscillations about the equilibrium point), with a constant amplitude A and a constant frequency ν (i.e., the number of cycles per second; see Figure 7.1). In addition to its amplitude and frequency, the motion of a simple harmonic oscillator is also characterized by a constant period T. This is the time for a single complete oscillation, with $T = 1/\nu$. The phase ϕ of the system determines the starting point on the sinusoidal wave.

In Eq. 7.12, the x-coordinate of the particle is oscillating back and forth between the A and $-A$ extremes around the rest position. Such a motion is found to correspond to the projection on the x-axis of the top of a vector with radius A, which is rotating at a constant speed in the xy-plane around the origin at $x = 0$ (Figure 7.2). This is a *phasor*. The speed $\omega = 2\pi\nu$ is also called the *angular velocity*, measured in units of radians per second; the oscillator frequency ν counts the number of full oscillations per second and is expressed in Hertz.

Using the techniques of differential calculus, the velocity v and acceleration a as a function of time t can be found:

$$v(t) = \frac{dx}{dt} = -A\omega\sin(\omega t + \phi); \tag{7.13}$$

$$a(t) = \frac{d^2x}{dt^2} = -A\omega^2\cos(\omega t + \phi). \tag{7.14}$$

The speed is maximal for $x = 0$ and drops to zero at the turning points, where the motion comes to a halt and then takes off again in the opposite direction. In this way, the particle spends most of its time in the extremal positions, and the least amount

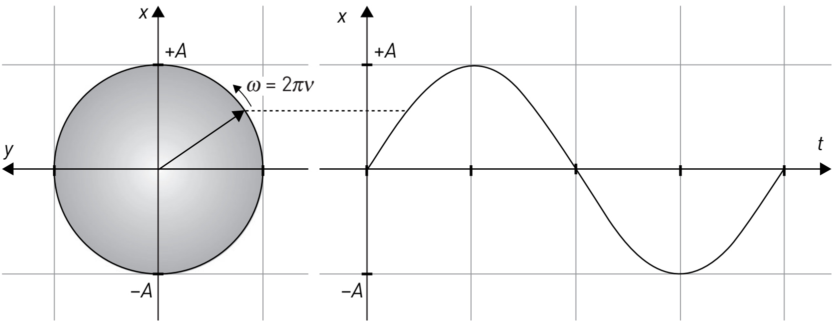


FIGURE 7.2 Representation of a sinusoidal function with the help of a *phase vector* or *phasor* that rotates with an angular velocity $\omega = 2\pi\nu$ about the origin in the xy-plane. The sine function is obtained by projecting the vector onto the x-axis. The amplitude A is the modulus of the vector.

of time in the equilibrium position. As we will see, this is different from the quantum mechanical ground state (see §7.2.2).

Using Eq. (7.12), the acceleration a in Eq. (7.14) can also be expressed as a function of the displacement x:

$$a(x) = -\omega^2 x. \tag{7.15}$$

Substituting Eq. (7.15) into Eq. (7.11) gives

$$\begin{aligned} ma &= m\left(-\omega^2 x\right) = -kx; \\ \omega &= \sqrt{\frac{k}{m}}. \end{aligned} \tag{7.16}$$

Since $\omega = 2\pi\nu$, the frequency ν of the oscillations is given by

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}}. \tag{7.17}$$

The period $T = 1/\nu$ denotes the time for one complete oscillation, and it is given by

$$T = 2\pi\sqrt{\frac{m}{k}}. \tag{7.18}$$

These equations demonstrate that the period T and frequency ν are independent of the amplitude A and the initial phase of motion ϕ.

Energy of a simple harmonic oscillator

Now let us consider the energy. The potential energy V is related to the displacement of the force:

$$F = -\frac{dV}{dx} = -kx, \tag{7.19}$$

where use was made of Eq. (7.10). Integrating Eq. (7.19) gives

$$\begin{aligned} \int dV &= \int kx dx; \\ V &= \frac{1}{2}kx^2 + C, \end{aligned} \tag{7.20}$$

where C is the integration constant. It follows from Eq. (7.16) that

$$k = m\omega^2. \tag{7.21}$$

Choosing the bottom of the potential well as the zero point of energy (i.e., $C = 0$), we have

$$V = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2 x^2. \tag{7.22}$$

This is the harmonic oscillator potential—one of the most important potentials in theoretical physics. Note that the graph of potential energy V versus the displacement x is a parabola. The vibrating spring is thus equivalent to a particle in a parabolic well. The kinetic energy T is

$$T = \frac{mv^2}{2} = \frac{(mv)^2}{2m} = \frac{p^2}{2m}. \tag{7.23}$$

The total energy E can be found by adding the potential energy V (Eq. (7.22)) to the kinetic energy T (Eq. (7.23)):

$$E = T + V = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{7.24}$$

This expression yields the Hamiltonian of the classical problem. Conservation of energy is guaranteed by

$$
\begin{aligned}
T &= \frac{mv^2}{2} = \frac{1}{2} mA^2\omega^2 \sin^2(\omega t + \phi);\\
V &= \frac{m\omega^2 x^2}{2} = \frac{1}{2} mA^2\omega^2 \cos^2(\omega t + \phi);\\
E &= T + V = \frac{mA^2\omega^2}{2} = \frac{kA^2}{2}.
\end{aligned} \tag{7.25}
$$

The total energy is thus only dependent on the square of the amplitude and the force constant of the elastic potential.

7.2.2 Quantum mechanical treatment

The *quantum harmonic oscillator* is the quantum mechanical analogue of the classical harmonic oscillator. It is one of the most important model systems in quantum mechanics because an arbitrary potential well can be approximated as a harmonic potential at the vicinity of a stable equilibrium point. This technique is of great importance for the description of molecular vibrations. Furthermore, the quantum harmonic oscillator is one of the few quantum mechanical systems for which a simple, exact solution is known. It is used as a basis for nuclear structure theory and quantum field theory, and its symmetry properties have been applied successfully in elementary particle physics, leading Murray Gell-Mann and Yuval Ne'eman to the *eightfold way*, as we shall see in §7.5.

Harmonic oscillator Hamiltonian

Analogous to the classical mechanical expression for the total energy (Eq. (7.24)), the harmonic oscillator Hamiltonian energy operator is

$$
\hat{\mathscr{H}} = \hat{T} + \hat{V} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}^2, \tag{7.26}
$$

where $\hat{x} = x$ is the position operator and $\hat{p}$ is the momentum operator, given by

$$
\hat{p} = -i\hbar \frac{d}{dx} = \frac{\hbar}{i}\frac{d}{dx}, \tag{7.27}
$$

with $\hbar = h/2\pi$ denoting Planck's constant—the logo of quantum mechanics (see Appendix D). Eq. (7.26) can be rewritten as follows:

$$
\hat{\mathscr{H}} = \frac{1}{2m}\left[\hat{p}^2 + m^2\omega^2\hat{x}^2\right]. \tag{7.28}
$$

To find the energy levels of the one-dimensional quantum mechanical harmonic oscillator and the corresponding energy eigenstates, we must solve the time-independent Schrödinger equation

$$
\hat{\mathscr{H}}\,|\psi\rangle = E\,|\psi\rangle. \tag{7.29}
$$

With Eqs. (7.27) and (7.28), this becomes

$$
\frac{1}{2m}\left(-\hbar^2\frac{d^2}{dx^2} + m^2\omega^2x^2\right)|\psi\rangle = E\,|\psi\rangle. \tag{7.30}
$$

We could solve this eigenvalue problem using the so-called *spectral method*, which consists of solving the differential equation (7.30), but it turns out that there is a different approach that circumvents the need to integrate and work with Hermite polynomials. In the following, we reapply the *ladder operator method* of Paul Dirac, which was described in Chapter 5 in relation to the angular momentum eigenstates. This alternative method is of an *algebraic* rather than an *analytic* nature; it doesn't involve calculus and therefore allows us to extract the energy eigenvalues E *without* having to solve the differential equation (7.30) directly. Moreover, it is applicable to more complicated problems, such as many-particle systems.

The operator method formalism

We start with a brief recapitulation of Dirac's ladder operator method. Let $|a\rangle$ be an eigenket of $\hat{A}$ with the eigenvalue equation

$$\hat{A}|a\rangle = a|a\rangle. \tag{7.31}$$

Operators $\hat{B}_+$ and $\hat{B}_-$ are said to be ladder operators for $\hat{A}$ when their respective commutation relations satisfy the following property:

$$\left[\hat{A}, \hat{B}_\pm\right] = \hat{A}\hat{B}_\pm - \hat{B}_\pm \hat{A} = \pm c \hat{B}_\pm, \tag{7.32}$$

for some nonzero real scalar c.

Factorization of the Hamiltonian

Following this approach, we can try to apply Dirac's method to the quantum mechanical harmonic oscillator problem. In this case, Eq. (7.31) is substituted with the time-independent Schrödinger equation, given in Eq. (7.29); operator $\hat{A}$ becomes the harmonic oscillator Hamiltonian $\hat{\mathscr{H}}$, and $\hat{B}_+$ and $\hat{B}_-$ take on the function of ladder operators, which should satisfy the following commutation relation:

$$\left[\hat{\mathscr{H}}, \hat{B}_\pm\right] = \hat{\mathscr{H}}\hat{B}_\pm - \hat{B}_\pm \hat{\mathscr{H}} = \pm c \hat{B}_\pm. \tag{7.33}$$

To find a valuable expression for the raising operator $\hat{B}_+$ and the lowering operator $\hat{B}_-$, it is helpful to note that the expression between square brackets in the harmonic oscillator Hamiltonian of Eq. (7.28) represents a sum of squares. This raises the possibility of *factorizing* the Hamiltonian—that is, decomposing it in a product of factors, which is a common operation within the field of algebra. Recall that a difference of squares $a^2 - b^2$ can be written as

$$a^2 - b^2 = (a+b)(a-b), \tag{7.34}$$

whereas a sum of squares $a^2 + b^2$ equals

$$a^2 + b^2 = (a+ib)(a-ib). \tag{7.35}$$

Based on this observation, we could attempt to express the Hamiltonian in Eq. (7.28) as a product of two factors:

$$\hat{\mathscr{H}} \overset{?}{=} \frac{1}{2m}\left(m\omega\hat{x} - i\hat{p}\right)\left(m\omega\hat{x} + i\hat{p}\right). \tag{7.36}$$

Let us check this factorization by calculating the product on the right-hand side of Eq. (7.36):

$$\begin{aligned}\frac{1}{2m}\left(m\omega\hat{x}-i\hat{p}\right)\left(m\omega\hat{x}+i\hat{p}\right)&=\frac{1}{2m}\left(\hat{p}^2+m^2\omega^2\hat{x}^2-im\omega\hat{p}\hat{x}+im\omega\hat{x}\hat{p}\right)\\&=\frac{1}{2m}\left(\hat{p}^2+m^2\omega^2\hat{x}^2\right)+\frac{i}{2}\omega\left(\hat{x}\hat{p}-\hat{p}\hat{x}\right)\\&=\hat{\mathscr{H}}+\frac{i}{2}\omega\left[\hat{x},\hat{p}\right],\end{aligned}\tag{7.37}$$

where we used Eq. (7.28) in the last line. If $\hat{x}$ and $\hat{p}$ were commuting with each other (i.e., $\left[\hat{x},\hat{p}\right]=0$), then Eq. (7.36) would be perfectly valid. However, using Eq. (7.27), the commutator $\left[\hat{x},\hat{p}\right]$ in Eq. (7.37) can be rewritten as

$$\left[\hat{x},\hat{p}\right]=\hat{x}\hat{p}-\hat{p}\hat{x}=\frac{\hbar}{i}\left(x\frac{d}{dx}-\frac{d}{dx}x\right).\tag{7.38}$$

Operating with Eq. (7.38) on some arbitrary function $f(x)$ gives

$$\begin{aligned}\frac{\hbar}{i}\left(x\frac{d}{dx}-\frac{d}{dx}x\right)f(x)&=\frac{\hbar}{i}\left[x\frac{d}{dx}f(x)-\frac{d}{dx}xf(x)\right]\\&=\frac{\hbar}{i}\left[x\frac{d}{dx}f(x)-\left(f(x)\frac{d}{dx}x+x\frac{d}{dx}f(x)\right)\right]\\&=\frac{\hbar}{i}\left[x\frac{d}{dx}f(x)-\left(1+x\frac{d}{dx}\right)f(x)\right]\\&=\frac{\hbar}{i}\left[x\frac{d}{dx}-1-x\frac{d}{dx}\right]f(x)\\&=-\frac{\hbar}{i}f(x).\end{aligned}\tag{7.39}$$

From this, follows the commutation relation

$$\left[\hat{x},\hat{p}\right]=-\frac{\hbar}{i},\tag{7.40}$$

and after substitution in Eq. (7.37), this yields

$$\frac{1}{2m}\left(m\omega\hat{x}-i\hat{p}\right)\left(m\omega\hat{x}+i\hat{p}\right)=\hat{\mathscr{H}}-\frac{1}{2}\hbar\omega.\tag{7.41}$$

Clearly, the proposed factorization in Eq. (7.36) is inadequate, because it differs from the Hamiltonian by the additive constant $-\hbar\omega/2$.

Also, although for numbers the products $(a+ib)(a-ib)$ and $(a-ib)(a+ib)$ are the same, for operators this depends on whether they commute. Hence, when we transcribe the classical Hamiltonian into the quantum one, we are once again facing the ambiguity that the order of the operators does matter. Let us examine this by reversing the order of the factors in Eq. (7.36) and calculating the product once again:

$$\begin{aligned}\frac{1}{2m}\left(m\omega\hat{x}+i\hat{p}\right)\left(m\omega\hat{x}-i\hat{p}\right)&=\frac{1}{2m}\left(\hat{p}^2+m^2\omega^2\hat{x}^2+im\omega\hat{p}\hat{x}-im\omega\hat{x}\hat{p}\right)\\&=\frac{1}{2m}\left(\hat{p}^2+m^2\omega^2\hat{x}^2\right)-\frac{i}{2}\omega\left(\hat{x}\hat{p}-\hat{p}\hat{x}\right)\\&=\hat{\mathscr{H}}-\frac{i}{2}\omega\left[\hat{x},\hat{p}\right]\\&=\hat{\mathscr{H}}+\frac{1}{2}\hbar\omega.\end{aligned}\tag{7.42}$$

Compared with Eq. (7.41), a different result is obtained because of the noncommutativity of $\hat{x}$ and $\hat{p}$. Once again, the factorization is seen to fail because it differs from the actual Hamiltonian by the additive constant $\hbar\omega/2$.

Note, however, that the sum of the two expressions in Eqs. (7.41) and (7.42) yields the Hamiltonian multiplied by 2:

$$\begin{aligned} &\frac{1}{2m}\left(m\omega\hat{x}-i\hat{p}\right)\left(m\omega\hat{x}+i\hat{p}\right)+\frac{1}{2m}\left(m\omega\hat{x}+i\hat{p}\right)\left(m\omega\hat{x}-i\hat{p}\right) \\ &\quad=\hat{\mathscr{H}}-\frac{1}{2}\hbar\omega+\hat{\mathscr{H}}+\frac{1}{2}\hbar\omega=2\hat{\mathscr{H}}. \end{aligned} \tag{7.43}$$

This allows us to construct the quantum Hamiltonian as the average of both possibilities. We start by defining

$$a^{\dagger}=\frac{1}{\sqrt{2m}}\left(m\omega\hat{x}-i\hat{p}\right) \tag{7.44}$$

as the *creation operator* (or the raising operator) and

$$a=\frac{1}{\sqrt{2m}}\left(m\omega\hat{x}+i\hat{p}\right) \tag{7.45}$$

as the *annihilation operator* (or the lowering operator) for reasons that will become obvious in the next section. As can be seen from Eqs. (7.44) and (7.45), the annihilation operator is the *adjoint* of the creation operator.[9] The Hamiltonian in Eq. (7.43) can then be written as

$$\hat{\mathscr{H}}=\frac{1}{2}\left(a^{\dagger}a+aa^{\dagger}\right). \tag{7.46}$$

Similarly, Eqs. (7.41) and (7.42) reduce to

$$\hat{\mathscr{H}}=a^{\dagger}a+\frac{1}{2}\hbar\omega; \tag{7.47}$$

$$\hat{\mathscr{H}}=aa^{\dagger}-\frac{1}{2}\hbar\omega. \tag{7.48}$$

With this, the commutation relation between $a^{\dagger}$ and a is found immediately:

$$\begin{aligned} \left[a^{\dagger},a\right]&=a^{\dagger}a-aa^{\dagger} \\ &=\hat{\mathscr{H}}-\frac{1}{2}\hbar\omega-\hat{\mathscr{H}}-\frac{1}{2}\hbar\omega \\ &=-\hbar\omega. \end{aligned} \tag{7.49}$$

Creation and annihilation operators

Before we continue, it should be noted that the two ladder operators $a^{\dagger}$ and a, together with the Hamiltonian $\hat{\mathscr{H}}$ and the identity operator $\hat{E}$, form a complete algebra. The corresponding symmetry is quite involved, however, and will only be discussed in Chapter 11. The rest of this section is devoted to the solution of the time-independent Schrödinger equation

$$\hat{\mathscr{H}}\left|\psi\right\rangle=E\left|\psi\right\rangle \tag{7.50}$$

---

[9] Note that the momentum operator is self-adjoint; see C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Mécanique Quantique.* Paris: Hermann, 1973.

for the one-dimensional harmonic oscillator, making use of the algebra of ladder operators $a^\dagger$ and a. Acting on $|\psi\rangle$ with the creation operator $a^\dagger$, we obtain

$$\begin{aligned}\hat{\mathscr{H}} a^\dagger |\psi\rangle &= \left(a^\dagger a + \frac{\hbar\omega}{2}\right) a^\dagger |\psi\rangle \\ &= a^\dagger a a^\dagger |\psi\rangle + \frac{\hbar\omega}{2} a^\dagger |\psi\rangle,\end{aligned} \tag{7.51}$$

where Eq. (7.47) was used. Since $aa^\dagger = a^\dagger a + \hbar\omega$ according to Eq. (7.49), Eq. (7.51) can be rewritten as follows:

$$\begin{aligned}\hat{\mathscr{H}} a^\dagger |\psi\rangle &= a^\dagger \left(a^\dagger a + \hbar\omega\right) |\psi\rangle + \frac{\hbar\omega}{2} a^\dagger |\psi\rangle \\ &= a^\dagger \left(a^\dagger a + \frac{\hbar\omega}{2}\right) |\psi\rangle + \hbar\omega a^\dagger |\psi\rangle \\ &= a^\dagger \hat{\mathscr{H}} |\psi\rangle + \hbar\omega a^\dagger |\psi\rangle \\ &= a^\dagger E |\psi\rangle + \hbar\omega a^\dagger |\psi\rangle \\ \hat{\mathscr{H}} a^\dagger |\psi\rangle &= (E + \hbar\omega)\, a^\dagger |\psi\rangle,\end{aligned} \tag{7.52}$$

where use was made again of Eq. (7.47). It thus follows from Eqs. (7.50) and (7.52) that if $|\psi\rangle$ is an eigenstate of $\hat{\mathscr{H}}$ with eigenvalue E, then $a^\dagger |\psi\rangle$ is an eigenfunction of $\hat{\mathscr{H}}$, with an eigenvalue $E + \hbar\omega$. In other words, operating on the eigenfunction $|\psi\rangle$ with the raising operator $a^\dagger$ converts $|\psi\rangle$ into another eigenfunction of $\hat{\mathscr{H}}$ with eigenvalue $\hbar\omega$ higher than the eigenvalue of $|\psi\rangle$.

If we now apply the raising operator $a^\dagger$ to Eq. (7.52) again, we find similarly

$$\hat{\mathscr{H}} \left(a^\dagger\right)^2 |\psi\rangle = (E + 2\hbar\omega) \left(a^\dagger\right)^2 |\psi\rangle. \tag{7.53}$$

Repeated application of the raising operator gives

$$\hat{\mathscr{H}} \left(a^\dagger\right)^k |\psi\rangle = (E + k\hbar\omega) \left(a^\dagger\right)^k |\psi\rangle, \qquad k = 0,\ 1,\ 2, \ldots. \tag{7.54}$$

In the same manner, the lowering operator a, when acting on the eigenstate $|\psi\rangle$, lowers the initial eigenvalue E with the amount $\hbar\omega$:

$$\begin{aligned}\hat{\mathscr{H}} a |\psi\rangle &= (E - \hbar\omega)\, a |\psi\rangle; \\ \hat{\mathscr{H}} a^k |\psi\rangle &= (E - k\hbar\omega)\, a^k |\psi\rangle.\end{aligned} \tag{7.55}$$

Thus, by using the raising and lowering operators on the eigenfunction $|\psi\rangle$ with the eigenvalue E, we generate a ladder of eigenvalues, the difference from step to step being $\hbar\omega$:

$$\ldots, \quad E - 2\hbar\omega, \quad E - \hbar\omega, \quad E, \quad E + \hbar\omega, \quad E + 2\hbar\omega, \quad \ldots. \tag{7.56}$$

We can now understand why the ladder operators for the quantum mechanical harmonic oscillator $a^\dagger$ and a were termed *creation* and *annihilation* operators, respectively. Indeed, because the raising operator $a^\dagger$ adds a quantum of energy $\hbar\omega$ to the oscillator system (Eq. (7.52)), thus *creating* a new energy shell, it can be called a *creation operator*. Similarly, the lowering operator a is interpreted as an annihilation operator because it subtracts a quantum of energy $\hbar\omega$ to the oscillator system, thus *annihilating* the initial energy shell (Eq. (7.55)).

The vacuum state

Given any energy eigenstate $|\psi\rangle$, we can act on it with the lowering operator a to produce another eigenstate $a|\psi\rangle$ with less energy. By repeated application of the lowering operator, it seems we can produce energy eigenstates down to $E=-\infty$. However, the set of eigenvalues of $\hat{\mathscr{H}}$ generated using the ladder operators $a^\dagger$ and a must have a *lower bound*, which can be seen as follows: for an eigenstate $|\psi\rangle$, the length of the square of the ket $a|\psi\rangle$ must be greater than or equal to zero, which implicates a lower bound on the energy:

$$E-\frac{\hbar\omega}{2}=\langle\psi|\,a^\dagger a\,|\psi\rangle=\langle a\psi|a\psi\rangle\geq 0, \tag{7.57}$$

which implies

$$E\geq\frac{\hbar\omega}{2}. \tag{7.58}$$

Therefore, let E_0 be the smallest eigenvalue of $\hat{\mathscr{H}}$ and let $|\psi_0\rangle$ be the corresponding eigenfunction. This eigenstate is often referred to as the *vacuum state* and must be destroyed when we act on it with the lowering operator:

$$a|\psi_0\rangle=0, \tag{7.59}$$

because otherwise $a|\psi_0\rangle$ would be an eigenfunction with eigenvalue $E_0-\hbar\omega<E_0$. We can see easily that $|\psi_0\rangle$ is an eigenstate of $\hat{\mathscr{H}}$ with the aid of the Schrödinger equation (Eq. (7.50)) and Eq. (7.47):

$$\begin{aligned}\hat{\mathscr{H}}\,|\psi_0\rangle&=\left[a^\dagger a+\frac{\hbar\omega}{2}\right]|\psi_0\rangle\\&=a^\dagger\,[a|\psi_0\rangle]+\frac{\hbar\omega}{2}\,|\psi_0\rangle\\&=0+\frac{\hbar\omega}{2}\,|\psi_0\rangle=E_0\,|\psi_0\rangle.\end{aligned} \tag{7.60}$$

It thus follows that the lowest achievable energy E_0 is exactly equal to $\hbar\omega/2$, which is called the *ground state energy* or *zero-point energy*. Classically, the oscillating particle could come to rest at the bottom of the potential well, where it has zero momentum p and a well-defined position x. Quantum mechanically, however, its energy cannot be zero.[10] According to Heisenberg's uncertainty principle $\Delta x\,\Delta p\geqslant\hbar/2$, the position x and momentum p of the oscillating particle cannot be known simultaneously with exact certainty (see Appendix D). This explains why the zero-point energy does not vanish. The particle is never at rest; it oscillates back and forth in the well. This view is confirmed when we look at the wave function for the ground state, which can be easily derived from Eq. (7.59). We obtain a first-order differential equation:

$$\begin{aligned}\frac{1}{\sqrt{2m}}\left[m\omega\hat{x}+i\hat{p}\right]|\psi_0\rangle&=0\\\hbar\frac{d}{dx}\,|\psi_0\rangle&=-m\omega x\,|\psi_0\rangle\\\frac{d\,|\psi_0\rangle}{|\psi_0\rangle}&=-\frac{m\omega}{\hbar}x\,dx.\end{aligned} \tag{7.61}$$

---

[10] If the particle comes to a complete rest, it ceases to have wavelike characteristics and becomes a classical particle.

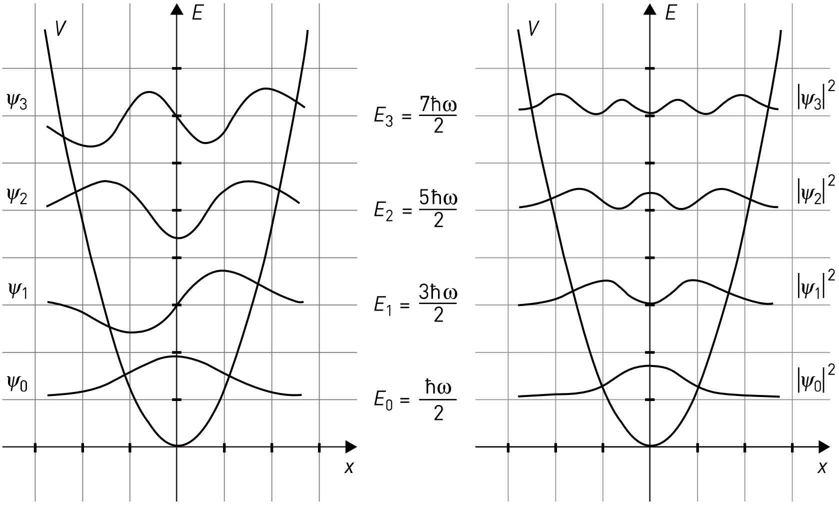


FIGURE 7.3 The wave functions $|\psi_i\rangle$, allowed energies E_i, and corresponding probability densities $|\psi_i|^2$ for the first four energy levels ($i = 0 \rightarrow 3$) of the quantum harmonic oscillator. The different wave functions are sinusoidal inside the potential well, but decay exponentially outside. The energy levels are equidistant.

This equation is easily solved. The normalized ground state wave function is given by

$$|\psi_0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right). \tag{7.62}$$

The ground state wave function thus has a bell shape, with its maximum at $x = 0$. Because the square of the wave function reflects the density distribution, it is implied that the particle is concentrated in a narrow zone around the equilibrium position. This contrasts with the weight distribution in the classical picture, where the density peaks at the extremal positions. For higher excited states of the oscillator, the wave function spreads out gradually to the extremal turning points, reaching a better correspondence between the classical and the quantum mechanical picture (Figure 7.3).

Excited states

To find the energy level of the nth eigenstate for the Hamiltonian, we use the fact that any eigenstate can be represented as the ground state $|\psi_0\rangle$ acted on n times by the raising operator: $\left(a^\dagger\right)^n|\psi_0\rangle$. Doing so, we find that the energy levels are

$$E_n = \frac{\hbar\omega}{2} + n\hbar\omega = \left(n + \frac{1}{2}\right)\hbar\omega, \tag{7.63}$$

with $n = 0, 1, 2, \ldots$ as the *vibration quantum number*. Interestingly, since $|\psi_0\rangle$ represents the vacuum state, it appears as if the whole energy spectrum has been created out of nothing (i.e., a quantum mechanical realization of the theologians' *creatio ex nihilo*, standing in sharp contrast with the famous Latin dictum *ex nihilo nihil fit*—from nothing, comes nothing). But, the energy spectrum of the quantum mechanical harmonic oscillator is noteworthy for a number of other reasons as well. First, the energies are *quantized* and may only take the discrete half-integer multiples of $\hbar\omega$. This is a typical feature of many quantum mechanical systems. Second, the energy levels are equally spaced (Figure 7.3), unlike the energy levels for the hydrogen atom (Chapter 9) or the particle in a box.

The symmetry of the one-dimensional oscillator

The Hamiltonian of the oscillator contains a product of two conjugate operators $a^\dagger a$. Such a form is invariant under a phase change of both operators. For $a' = \exp(-i\theta)\,a$, the adjoint operator transforms as the complex conjugate, $(a^\dagger)' = a^\dagger \exp(i\theta)$, so that the bilinear form remains unchanged:

$$(a^\dagger)' a' = a^\dagger \exp(i\theta) \exp(-i\theta)\,a = a^\dagger a. \tag{7.64}$$

Clearly, the set of all phase factors forms a group, the elements of which are characterized by an angle θ in the complex plane. The corresponding group is called the *unitary group* in one dimension: U(1). We will come back to the U(1) group in §7.4.3. For the moment, it suffices to say that a change of phase factor can be viewed as a rotation on the unit circle in the complex plane. As a result, U(1) is similar to the two-dimensional rotation group SO(2). From a shell theoretical perspective, U(1) does not contain additional information because all levels are nondegenerate. In contrast, in higher dimensional cases, degeneracies arise and the unitary groups act as true degeneracy groups, as we shall see in a moment.

The algebra of the ladder operators is a different matter. The ladder operators do not conserve the energy; like elevators, they run up and down the oscillator levels. A symmetry group that hosts the ladder operators thus collects all the levels in one representational set. Such an algebra is called a *spectrum generating algebra* because it changes the energy as it moves through the set of levels. We will come back to this concept when addressing the noncompact Lie group SO(2,1) in Chapter 11.

7.3 THE THREE-DIMENSIONAL HARMONIC OSCILLATOR

The one-dimensional harmonic oscillator is easily generalizable to n dimensions. In one dimension, the position of the particle is specified by a single x coordinate. In n dimensions, this is replaced by n Cartesian position coordinates labeled $x_1, x_2, \ldots, x_n$. Corresponding to each position coordinate is a momentum; these are labeled $p_1, p_2, \ldots, p_n$. In analogy with the one-dimensional Hamiltonian, given by Eq. (7.26), the Hamiltonian for the n-dimensional quantum harmonic oscillator is

$$\hat{\mathscr{H}} = \sum_{i=1}^{n} \left(\frac{\hat{p}_i^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}_i^2 \right), \tag{7.65}$$

where $\hat{x}_i$ is the ith position operator, and $\hat{p}_i$ is the ith momentum operator. As the form of this Hamiltonian makes clear, the n-dimensional harmonic oscillator is exactly analogous to n independent one-dimensional harmonic oscillators with the same mass m and spring constant ω. In this case, however, each of the quantities $x_1, x_2, \ldots, x_n$ refers to the position of one of the n particles.

In the specific case of a spherically symmetric three-dimensional harmonic oscillator, the three position coordinates x_1, x_2, and x_3 can be written as x, y, and z.

Similarly, the momenta p_1, p_2, and p_3 become p_x, p_y, and p_z, respectively. Following Eq. (7.65), the Hamiltonian energy operator then equals

$$\begin{aligned}\hat{\mathscr{H}} &= \hat{\mathscr{H}}_x + \hat{\mathscr{H}}_y + \hat{\mathscr{H}}_z \\ &= \frac{\hat{p}_x^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 + \frac{\hat{p}_y^2}{2m} + \frac{1}{2}m\omega^2\hat{y}^2 + \frac{\hat{p}_z^2}{2m} + \frac{1}{2}m\omega^2\hat{z}^2 \\ &= \frac{1}{2m}\left[\hat{p}_x^2 + m^2\omega^2\hat{x}^2\right] + \frac{1}{2m}\left[\hat{p}_y^2 + m^2\omega^2\hat{y}^2\right] + \frac{1}{2m}\left[\hat{p}_z^2 + m^2\omega^2\hat{z}^2\right].\end{aligned} \tag{7.66}$$

This system consists of a particle with mass m bound to the origin by a force $\mathbf{F}$ whose components F_x, F_y, and F_z along the x-, y-, and z-axes are equal to $-k_x x$, $-k_y y$, and $-k_z z$, respectively. k_x, k_y, and k_z are the three force constants in the three directions, and x, y, and z are the components of the displacement along the three axes. In the case of spherical symmetry, the spring constant has the same magnitude in all directions: $k_x = k_y = k_z$. This oscillator is often referred to as an *isotropic oscillator*. Following the same ladder operator approach as used with the one-dimensional harmonic oscillator, the Hamiltonian $\hat{\mathscr{H}}$ can be rewritten in analogy with Eq. (7.47) as

$$\begin{aligned}\hat{\mathscr{H}} &= \left(a_x^\dagger a_x + \frac{\hbar\omega}{2}\right) + \left(a_y^\dagger a_y + \frac{\hbar\omega}{2}\right) + \left(a_z^\dagger a_z + \frac{\hbar\omega}{2}\right) \\ &= a_x^\dagger a_x + a_y^\dagger a_y + a_z^\dagger a_z + \frac{3}{2}\hbar\omega.\end{aligned} \tag{7.67}$$

Note that the form of this expression is analogous to the one-dimensional expression for the Hamiltonian in Eq. (7.47). As a result of the independence of the three spatial directions, the three-dimensional Hamiltonian $\hat{\mathscr{H}}$ can be seen as the sum of three Hamiltonians: $\hat{\mathscr{H}}_x$ acting in E_x, $\hat{\mathscr{H}}_y$ acting in E_y, and $\hat{\mathscr{H}}_z$ acting in E_z. Each of these Hamiltonians relates to a one-dimensional Schrödinger equation:

$$\hat{\mathscr{H}}_x\left|\psi_{n_x}\right\rangle = E_x\left|\psi_{n_x}\right\rangle = \left(n_x + \frac{1}{2}\right)\hbar\omega\left|\psi_{n_x}\right\rangle; \tag{7.68}$$

$$\hat{\mathscr{H}}_y\left|\psi_{n_y}\right\rangle = E_y\left|\psi_{n_y}\right\rangle = \left(n_y + \frac{1}{2}\right)\hbar\omega\left|\psi_{n_y}\right\rangle; \tag{7.69}$$

$$\hat{\mathscr{H}}_z\left|\psi_{n_z}\right\rangle = E_z\left|\psi_{n_z}\right\rangle = \left(n_z + \frac{1}{2}\right)\hbar\omega\left|\psi_{n_z}\right\rangle; \tag{7.70}$$

analogous to Eqs. (7.50) and (7.63). Since $\hat{\mathscr{H}} = \hat{\mathscr{H}}_x + \hat{\mathscr{H}}_y + \hat{\mathscr{H}}_z$,

$$\hat{\mathscr{H}}\left|\psi_{n_x,n_y,n_z}\right\rangle = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega\left|\psi_{n_x,n_y,n_z}\right\rangle = E_n\left|\psi_{n_x,n_y,n_z}\right\rangle. \tag{7.71}$$

The energy levels of the three-dimensional harmonic oscillator are thus denoted by $E_n = \left(n + 3/2\right)\hbar\omega$, with n as the total quantum number, being a nonnegative integer $n = n_x + n_y + n_z$. Because the energy of this system depends only on the sum of the quantum numbers n_x, n_y, and n_z, all the energy levels for the isotropic oscillator are degenerate, except for the lowest level $E_0 = 3/2\hbar\omega$. Indeed, for the first energy level, $n = 0$, there is only a single eigenstate: $|\psi_{000}\rangle$. The next level, $n = 1$, has three states—namely, $|\psi_{100}\rangle$, $|\psi_{010}\rangle$, and $|\psi_{001}\rangle$—so it is degenerate. A little further thought makes clear that higher energy states have even greater degeneracy (i.e., there are even more eigenstates sharing the same eigenvalue). The second energy level, $n = 2$, for instance, is sixfold degenerate with $|\psi_{200}\rangle$, $|\psi_{020}\rangle$, $|\psi_{002}\rangle$, $|\psi_{110}\rangle$, $|\psi_{101}\rangle$, and $|\psi_{011}\rangle$, all sharing the same eigenvalue $E_2 = 7/2\hbar\omega$. The degeneracy at a particular level n can be found by a smart rearrangement of the (n_x, n_y, n_z) quantum levels as follows:

$$\begin{aligned}\sum_{n_x+n_y+n_z=n} \left(n_x, n_y, n_z\right) = {} & [(n,0,0)] \\ & +[(n-1,1,0)+(n-1,0,1)] \\ & +[(n-2,2,0)+(n-2,1,1)+(n-2,0,2)] \\ & +\ldots \\ & +[(n-i,i,0)+(n-i,i-1,1)+\ldots(n-i,0,i)] \\ & +\ldots \\ & +[(0,n,0)+(0,n-1,1)+\ldots(0,0,n)].\end{aligned} \tag{7.72}$$

Here, the excitation is presented in slowly decreasing order of n_x. For $n_x = n-i$, there are $i+1$ couples of n_y, n_z. The total degeneracy D can thus be easily calculated:

$$D = \sum_{i=0}^{n}(i+1) = \frac{(n+1)(n+2)}{2}, \tag{7.73}$$

resulting in the sequence 1, 3, 6, 10, 15 for $n = 0 \rightarrow 4$.

Harmonic shells and atomic clusters

By way of illustration, let us briefly consider the shell structure of atomic clusters. An *atomic cluster* is a small multiatomic particle, usually of nanometer size. As such, it is an intermediate stage between the free atom and the regular solid. Clusters can be formed by firing an intense laser beam at a solid sample. The collision leads to the formation of a hot plasma of ionized atoms. When the plasma is cooled rapidly by the supersonic expansion of the helium carrier gas in a high vacuum, atoms aggregate to clusters that can be detected in a mass spectrometer.

Figure 7.4 shows the outcome of such an experiment in the case of sodium clusters.[11] It is clear that not all cluster sizes, or *nuclearities*, as they are called, are being formed with equal probability. The peaks in the spectrum indicate clusters with extra stability. The magic numbers associated with these are identified as 8, 20, and 40, and less pronounced magic features are seen at 58 and 92.

The simplest shell model to interpret these spectral results is based on the harmonic oscillator. In this model, a cluster of sodium atoms is considered as a kind of superatom. The sodium atoms have a lonely valence electron that is easily ionized. The resulting electron gas holds the Na^+ ions together in a spherical aggregate, giving rise to an effective central attraction potential comparable with a spherical parabolic well. The states of this superatom thus correspond to a first approximation to the energy levels of the harmonic oscillator. According to the Pauli principle for an electron system, each eigenstate can accommodate only two electrons with opposite spin. The

---

[11] W. D. Knight, K. Clemenger, W. A. de Heer, W. A. Saunders, M. Y. Chou, and M. L. Cohen. "Electronic Shell Structure and Abundances of Sodium Clusters." *Physical Review Letters* 52.24 (1984), pp. 2141–2143.

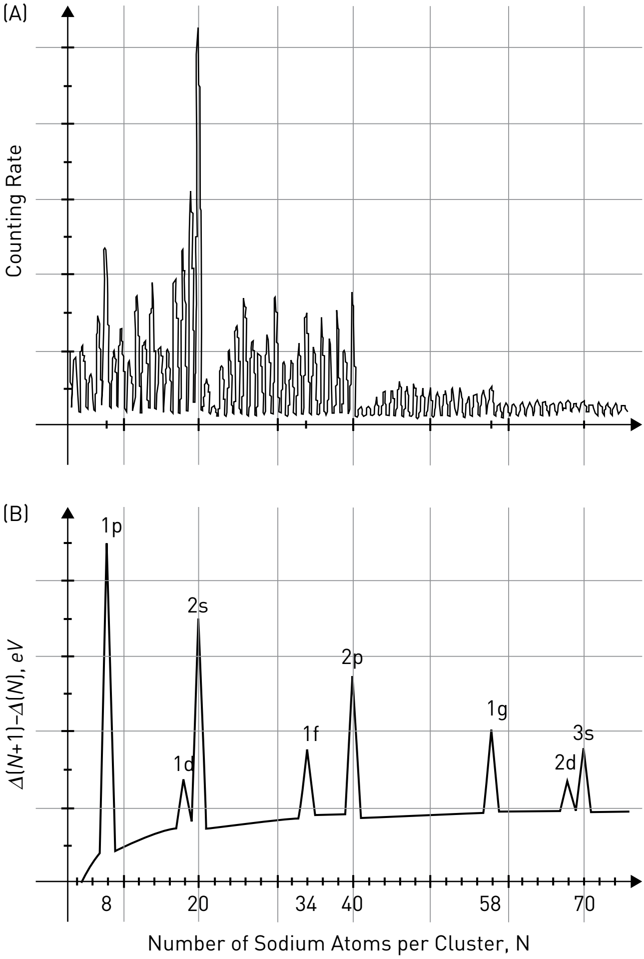


FIGURE 7.4 (A) Mass spectrum of sodium clusters, with $N = 4$ to 75. (B) Simulated shell model showing peaks corresponding to closed-shell states. Adapted from W. D. Knight, K. Clemenger, W. A. de Heer, W. A. Saunders, M. Y. Chou, and M. L. Cohen. “Electronic Shell Structure and Abundances of Sodium Clusters.” (1984).

number of electrons per shell therefore equals 2, 6, 12, 20, 30, leading to the following *magic numbers* for shell closure:

$$2, (2+6), (2+6+12), (2+6+12+20), (2+6+12+20+30), \tag{7.74}$$

or

$$2, 8, 20, 40, \text{ and } 70. \tag{7.75}$$

The numbers 8, 20, and 40 are indeed observed in the experiment. For higher nuclearities, deviations of the shell model become more important and lead to a breakdown of the oscillator shells in spherical subshells,[12] explaining the weak features at 58 and 92.

---

[12] The spherical subshells of a given oscillator shell are derived in Chapter 14.2.3.

The U(3) group and the isotropic harmonic oscillator

Until now, no mention has been made of the use of *symmetry groups*, but it would be surprising if the magic sequence of degeneracies would not hide a valiant degeneracy group.[13] According to D. M. Fradkin, "an important starting point in the investigation of a physical system is the determination of the *symmetries* it possesses. Once these are known, many of the system's properties can be established by quite general means. The mathematical description of a symmetry is in terms of a corresponding set of operations that leaves the system unchanged. By this we mean that the result of these operations still gives the same system although perhaps in a different state."[14] An example might help to understand this point. Suppose, for instance, that a certain operator $\hat{A}$ exists that acts on the eigenfunction $|\psi_{210}\rangle$, transforming it into the eigenstate $|\psi_{201}\rangle$. Because both functions belong to the $n = 3$ shell, the energy remains unchanged, although the transformed state $|\psi_{201}\rangle$ is clearly different from the initial eigenstate $|\psi_{210}\rangle$.

Let us verify whether any of the six previously mentioned ladder operators ($a_x^\dagger$, a_x, $a_y^\dagger$, a_y, $a_z^\dagger$, or a_z) is able to transform the eigenstate $|\psi_{210}\rangle$ into the eigenfunction $|\psi_{201}\rangle$. Clearly, such a transformation in only *one* step is impossible when use is made of these six ladder operators. At least *two* steps are needed. For instance, acting first with a_y on $|\psi_{210}\rangle$ lowers the eigenfunction, resulting in $|\psi_{200}\rangle$. In a second step, $a_z^\dagger$ can be used to raise $|\psi_{200}\rangle$ to $|\psi_{201}\rangle$. It thus follows that $a_z^\dagger a_y$ is a suitable ladder operator to transform $|\psi_{210}\rangle$ into $|\psi_{201}\rangle$ in just one step.

How many of these ladder operators could we possibly build? Because the general form of this type of ladder operator can be expressed as $a_i^\dagger a_j$, with $i = x, y, z$ and $j = x, y, z$, a little further consideration brings us to the conclusion that a total of *nine* such operators can be built: $a_x^\dagger a_x$, $a_x^\dagger a_y$, $a_x^\dagger a_z$, $a_y^\dagger a_x$, $a_y^\dagger a_y$, $a_y^\dagger a_z$, $a_z^\dagger a_x$, $a_z^\dagger a_y$, and $a_z^\dagger a_z$. The Hamiltonian itself is proportional to the sum of the three diagonal elements in this row (i.e. the elements with $i = j$; see Eq. (7.67)). This sum clearly corresponds to the Casimir operator, which characterizes the shell by the quantum number n. Once the operators are in place, the next task is to derive their commutation relations. From there, we obtain the algebra, and hence the symmetry group, and all it entails.

Instead of continuing this heuristic derivation, let us take a stepwise approach so as to discover the landscape gradually while we climb our way to the top, starting from the by-now-familiar valley of the orthogonal matrix group. In this way, the connection between SO(3) and SU(3) can be clearly exposed.

We begin by rewriting the three creation operators as a row vector $\mathbf{a}^\dagger$, and the three destruction operators as a column vector $\mathbf{a}$. The scalar product of these vectors corresponds to the operator part of the Hamiltonian:

$$\begin{aligned}\mathbf{a}^\dagger \cdot \mathbf{a} &= \begin{bmatrix} a_x^\dagger & a_y^\dagger & a_z^\dagger \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \\ &= a_x^\dagger a_x + a_y^\dagger a_y + a_z^\dagger a_z.\end{aligned} \tag{7.76}$$

---

[13] See H. J. Lipkin. "The Three-Dimensional Harmonic Oscillator." In: *Lie Groups for Pedestrians*. Mineola, NY: Dover Publications, 2002, pp. 57–68.

[14] D. M. Fradkin. "Three-Dimensional Isotropic Harmonic Oscillator and SU3." *American Journal of Physics* 33.3 (1965), pp. 207–211, 207.

This scalar product is very similar to the product form from which the rotation group was derived, with one difference though; the current vectors are no longer real, but complex entities. This means that a transformation of $\mathbf{a}^\dagger$ by a matrix $\mathbb{U}$, as $(\mathbf{a}^\dagger)' = \mathbf{a}^\dagger\,\mathbb{U}$, will be matched by a conjugate transformation of the column vector of annihilation operators (i.e., with a transformation matrix that is the transposed *and* complex conjugate of $\mathbb{U}$—transposed because row and columns are interchanged, complex conjugate because creation and annihilation operators are complex conjugate). We denote this conjugate matrix as $\mathbb{U}^\dagger$, where the dagger stands for transposition and complex conjugation: $\mathbb{U}^\dagger = (\mathbb{U}^*)^{\mathrm{T}}$. The scalar product is thus transformed as follows:

$$(\mathbf{a}^\dagger)' \cdot \mathbf{a}' = \mathbf{a}^\dagger\,\mathbb{U}\mathbb{U}^\dagger\,\mathbf{a}. \tag{7.77}$$

Conservation of this scalar product requires that the transformation matrices have the following property:

$$\mathbb{U}^\dagger\mathbb{U} = \mathbb{I} = \mathbb{U}\mathbb{U}^\dagger. \tag{7.78}$$

Matrices satisfying this property are said to be *unitary* (see Appendix B). Because $\mathbf{a}$ and $\mathbf{a}^\dagger$ are 3-vectors, the transformation matrices $\mathbb{U}$ and $\mathbb{U}^\dagger$ are of dimension 3. The set of all 3 × 3 unitary matrices forms a group, the group of unitary, unimodular, three-dimensional, linear transformations, U(3). The three-dimensional isotropic harmonic oscillator is therefore said to possess the U(3) group as a symmetry group. We will study the U(3) and SU(3) groups in detail in the next section.[15]

7.4 THE UNITARY GROUPS U(3) AND SU(3)

7.4.1 The unitary group U(3)

The unitary group U(3) and its special unitary subgroup SU(3) are of ubiquitous importance in physics. We will focus on the derivation of the generators and the Cartan-Weyl structure.

Definition 7.1 (The U(3) Lie group): The set of all 3 × 3 unitary matrices,

$$\mathrm{U}(3) = \left\{ \text{complex } 3 \times 3 \text{ matrices } \mathbb{U} : \begin{array}{l} \mathbb{U}^\dagger\mathbb{U} = \mathbb{U}\mathbb{U}^\dagger = \mathbb{I} \\ |\det \mathbb{U}| = 1 \end{array} \right\}, \tag{7.79}$$

---

[15] Other introductory accounts on the unitary groups can be found in J. F. Cornwell. *Group Theory in Physics: An Introduction*. San Diego: Academic Press, 1997, pp. 255–270; J. Fuchs and C. Schweigert. "The Lie Algebra $\mathfrak{su}(3)$ and Hadron Symmetries." In: *Symmetries, Lie Algebras and Representations: A Graduate Course for Physicists*. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press, 1997, pp. 30–46; S. Gasiorowicz. *Elementary Particle Physics*. New York: John Wiley & Sons, 1966; W. M. Gibson and B. R. Pollard. *Symmetry Principles in Elementary Particle Physics*. Cambridge: Cambridge University Press, 1976; W. Greiner and B. Müller. "The SU(3) Symmetry." In: *Quantum Mechanics: Symmetries*. Berlin: Springer-Verlag, 2001, pp. 195–229; W. Greiner and B. Müller. "Quarks and SU(3)." In: *Quantum Mechanics: Symmetries*. Berlin: Springer-Verlag, 2001, pp. 231–307; B. C. Hall. "The Representations of SU(3)." In: *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. New York: Springer-Verlag, 2010, pp. 127–153; H. F. Jones. "'Accidental' Degeneracy of the H Atom and SO(4)." In: *Groups, Representations and Physics*. New York: Taylor & Francis 1998, pp. 124–127; A. W. Joshi. "Dynamical Symmetry." In: *Elements of Group Theory for Physicists*. New Delhi: Wiley Eastern, 1977, pp. 171–176; H. J. Lipkin. "The Group SU(3) and Its Application to Elementary Particles." In: *Lie Groups for Pedestrians*. Mineola, NY: Dover Publications, 2002, pp. 33–56; W. Pfeifer. *The Lie Algebras su(N): An Introduction*. Basel, Switzerland: Birkhäuser, 2003.

forms a continuous, connected, compact Lie group under matrix multiplication. This group is called the *unitary group* in three dimensions and is denoted by the symbol U(3). ▪

Proof. Let $\mathbb{U}_a$ and $\mathbb{U}_b$ denote two 3×3 unitary matrices of the group U(3), and suppose that $\mathbb{U}_c = \mathbb{U}_a \mathbb{U}_b$.

1. *Closure*: We need to prove that $\mathbb{U}_c \in \mathrm{U}(3)$; (i.e., $\mathbb{U}_c$ is a 3×3 unitary matrix with a determinant of absolute value 1). Since $\mathbb{U}_a$ and $\mathbb{U}_b$ are 3×3 matrices, $\mathbb{U}_c$ is also a 3×3 matrix. Also,

$$\mathbb{U}_c^\dagger \mathbb{U}_c = (\mathbb{U}_a \mathbb{U}_b)^\dagger \mathbb{U}_a \mathbb{U}_b = \mathbb{U}_b^\dagger \mathbb{U}_a^\dagger \mathbb{U}_a \mathbb{U}_b = \mathbb{U}_b^\dagger \mathbb{U}_b = \mathbb{I} \text{ (unitarity)}; \tag{7.80}$$

$$|\det \mathbb{U}_c| = |\det(\mathbb{U}_a \mathbb{U}_b)| = |\det \mathbb{U}_a|\,|\det \mathbb{U}_b| = 1. \tag{7.81}$$

2. *Associativity*: Matrix multiplication is associative, so the associative law holds true for the U(3) group elements.
3. *Identity element*: The 3×3 identity matrix $\mathbb{I}$ represents the identity element.
4. *Inverse element*: Let $\mathbb{U}_a^{-1}$ denote the inverse matrix of $\mathbb{U}_a$. We need to prove that $\mathbb{U}_a^{-1} \in \mathrm{U}(3)$. Since $(\mathbb{U}_a^{-1})^\dagger = (\mathbb{U}_a^\dagger)^{-1}$,

$$\left(\mathbb{U}_a^{-1}\right)^\dagger \mathbb{U}_a^{-1} = \left(\mathbb{U}_a^\dagger\right)^{-1} \mathbb{U}_a^{-1} = \left(\mathbb{U}_a \mathbb{U}_a^\dagger\right)^{-1} = \mathbb{I}^{-1} = \mathbb{I} \text{ (unitarity)}; \tag{7.82}$$

$$\left|\det \mathbb{U}_a^{-1}\right| = \left|\det \mathbb{U}_a^\dagger\right| = \left|\det \mathbb{U}_a^*\right| = |\det \mathbb{U}_a| = 1. \tag{7.83}$$

This proves that U(3) forms a Lie group. ▪

7.4.2 Subgroups of the unitary group U(3)

The relationship between SU(3) and U(3) is analogous to the relationship between SO(3) and O(3).

Definition 7.2 (The SU(3) subgroup): The subset of 3×3 unitary matrices with *unit determinant*

$$\mathrm{SU}(3) = \left\{ \text{complex } 3 \times 3 \text{ matrices } \mathbb{U} : \begin{array}{l} \mathbb{U}^\dagger \mathbb{U} = \mathbb{U}\mathbb{U}^\dagger = \mathbb{I} \\ \det \mathbb{U} = 1 \end{array} \right\} \tag{7.84}$$

forms a subgroup of the unitary group U(3) under matrix multiplication. This group is called the *special unitary group* in three dimensions and is denoted by the symbol SU(3). ▪

Definition 7.3 (The O(3) and SO(3) subgroups): Following Eq. (5.18), the set of 3×3 orthogonal matrices with unimodular determinant

$$\mathrm{O}(3) = \left\{ \text{real } 3 \times 3 \text{ matrices } \mathbb{R} : \begin{array}{l} \mathbb{R}^\mathrm{T} \mathbb{R} = \mathbb{R}\mathbb{R}^\mathrm{T} = \mathbb{I} \\ |\det \mathbb{R}| = 1 \end{array} \right\} \tag{7.85}$$

forms a subgroup of the unitary group U(3) under matrix multiplication. This is the familiar *orthogonal group* in three dimensions, denoted by the symbol O(3). The subset of orthogonal matrices in Eq. (7.85) with unit determinant ($\det \mathbb{R} = 1$) forms the group SO(3), which is also a subgroup of the unitary group U(3). ▪

We denote the relationship between U(3) and SU(3) by a descending chain of subgroups:

$$\mathrm{U}(3) \supset \mathrm{SU}(3). \tag{7.86}$$

The same notation applies for the chain

$$\mathrm{U}(3) \supset \mathrm{O}(3) \supset \mathrm{SO}(3). \tag{7.87}$$

7.4.3 The generators of U(3)

Following the Lie group strategy, we must now determine the infinitesimal generators of U(3), which we denote by the symbols $\mathbb{X}_i$. To determine the generators $\mathbb{X}_i$, it will prove helpful to express the unitary matrices $\mathbb{U}$ in the form

$$\mathbb{U} = e^{i\mathbb{X}}, \tag{7.88}$$

with $\mathbb{X}$ a 3×3 square *Hermitian* matrix.

Proof of the Hermiticity of $\mathbb{X}$

To demonstrate the Hermiticity of $\mathbb{X}$ (i.e., $\mathbb{X}^\dagger = \mathbb{X}$), we start by noting that the inverse matrix $\mathbb{U}^{-1}$ equals

$$\mathbb{U}^{-1} = e^{-i\mathbb{X}}. \tag{7.89}$$

After all,

$$\mathbb{U}\mathbb{U}^{-1} = e^{i\mathbb{X}} e^{-i\mathbb{X}} = e^{i\mathbb{X} - i\mathbb{X}} = e^{\mathbb{O}} = \mathbb{I}. \tag{7.90}$$

As a result of the unitarity of the matrix $\mathbb{U}$ (cf. Eq. (7.78)), the inverse of the matrix $\mathbb{U}$ can also be written as $\mathbb{U}^{-1} = \mathbb{U}^\dagger$. Eq. (7.89) can then be rewritten as

$$\mathbb{U}^{-1} = \mathbb{U}^\dagger = \left[e^{i\mathbb{X}}\right]^\dagger = e^{(i\mathbb{X})^\dagger} = e^{-i\mathbb{X}^\dagger}, \tag{7.91}$$

where we have used the property that the adjoint of a matrix function equals $\left[f(\mathbb{A})\right]^\dagger = f(\mathbb{A}^\dagger)$. Because of the uniqueness of the inverse of the matrix $\mathbb{U}$, it follows that

$$e^{-i\mathbb{X}} = e^{-i\mathbb{X}^\dagger}, \tag{7.92}$$

which holds for

$$\mathbb{X} = \mathbb{X}^\dagger. \tag{7.93}$$

That is, $\mathbb{X}$ is a Hermitian matrix; the Hermitian adjoint $\mathbb{X}^\dagger$ is identical to $\mathbb{X}$. In terms of its matrix elements,

$$\begin{aligned} x_{ij} &= x_{ji}^* ; \\ R_{ij} + iI_{ij} &= R_{ji} - iI_{ji}. \end{aligned} \tag{7.94}$$

For the diagonal elements of $\mathbb{X}$, this property leads to

$$\begin{aligned} x_{ii} &= x_{ii}^* ; \\ R_{ii} + iI_{ii} &= R_{ii} - iI_{ii}, \end{aligned} \tag{7.95}$$

which holds only for $I_{ii} = 0$; that is, the diagonal elements of a Hermitian matrix are all *real*. Before we can continue our determination of the SU(3) generators, we need to acquaint ourselves with two more properties of Hermitian matrices: their *trace* and their number of independent matrix elements.

The trace of a Hermitian matrix

Recall the definition of the trace of an $n \times n$ matrix $\mathbb{A}$, denoted $\mathrm{Tr}(\mathbb{A})$, as the sum of the elements on the main diagonal

$$\mathrm{Tr}(\mathbb{A}) \equiv \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn} = a. \tag{7.96}$$

Eq. (7.95) then implies that the *trace* of the Hermitian matrix $\mathbb{X}$ is a real number R:

$$\mathrm{Tr}(\mathbb{X}) = \sum_{i=1}^{n} R_{ii} = R_{11} + R_{22} + \ldots + R_{nn} = R, \tag{7.97}$$

with R_{ii} the matrix element of $\mathbb{X}$ on the ith row and ith column.

The independent matrix elements of a Hermitian matrix

Following Eq. (7.94), we can express the general form of an $n \times n$ Hermitian matrix $\mathbb{X}$ as follows:

$$\mathbb{X} = \begin{bmatrix} R_{11} & R_{12} - iI_{12} & R_{13} - iI_{13} & \cdots & R_{1n} - iI_{1n} \\ R_{12} + iI_{12} & R_{22} & R_{23} - iI_{23} & \cdots & R_{2n} - iI_{2n} \\ R_{13} + iI_{13} & R_{23} + iI_{23} & R_{33} & \cdots & R_{3n} - iI_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{1n} + iI_{1n} & R_{2n} + iI_{2n} & R_{3n} + iI_{3n} & \cdots & R_{nn} \end{bmatrix}. \tag{7.98}$$

A general $n \times n$ complex matrix has $2n^2$ matrix elements, n^2 of which are real and another n^2 of which are imaginary. However, because of the Hermiticity of $\mathbb{X}$, the number of *independent* matrix elements is reduced. Let us first consider the number of real matrix elements of $\mathbb{X}$. The n diagonal elements $\mathbb{X}_{ii}$ are all real, leaving $n^2 - n$ triangular matrix elements. Because of Eq. (7.94), the upper triangular part of $\mathbb{X}$ is perfectly defined by the lower triangular part, leaving $\left(n^2 - n\right)/2$ unique triangular real matrix elements. The total number N_R of independent real matrix elements is then

$$N_R = n + \frac{n^2 - n}{2} = \frac{2n + n^2 - n}{2} = \frac{n^2 + n}{2}. \tag{7.99}$$

The same applies to the imaginary elements of $\mathbb{X}$, except that there are no imaginary elements on the diagonal. The total number N_I of unique imaginary elements is therefore given by

$$N_I = \frac{n^2 - n}{2}. \tag{7.100}$$

This leads to a simple formula that describes the total number N of independent matrix elements of an $n \times n$ Hermitian matrix $\mathbb{X}$:

$$N = N_R + N_I = \frac{n^2 + n}{2} + \frac{n^2 - n}{2} = \frac{2n^2}{2} = n^2. \tag{7.101}$$

The generators of U(3)

We start from the general Hermitian matrix $\mathbb{X}$ of order 3:

$$\mathbb{X} = \begin{bmatrix} R_{11} & R_{12} - iI_{12} & R_{13} - iI_{13} \\ R_{12} + iI_{12} & R_{22} & R_{23} - iI_{23} \\ R_{13} + iI_{13} & R_{23} + iI_{23} & R_{33} \end{bmatrix}. \tag{7.102}$$

Taking the partial derivative of $\mathbb{X}$ with respect to every unique matrix element R_{ij} or I_{ij} leads to $3^2 = 9$ linearly independent Hermitian matrices of order 3:

$$\begin{aligned}
&\mathbb{X}_1 = \tfrac{\partial}{\partial R_{12}}\mathbb{X} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \mathbb{X}_2 = \tfrac{\partial}{\partial I_{12}}\mathbb{X} = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};\\
&\mathbb{X}_3 = \tfrac{\partial}{\partial R_{11}}\mathbb{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \mathbb{X}_4 = \tfrac{\partial}{\partial R_{13}}\mathbb{X} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix};\\
&\mathbb{X}_5 = \tfrac{\partial}{\partial I_{13}}\mathbb{X} = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}; \quad \mathbb{X}_6 = \tfrac{\partial}{\partial R_{23}}\mathbb{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix};\\
&\mathbb{X}_7 = \tfrac{\partial}{\partial I_{23}}\mathbb{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}; \quad \mathbb{X}_8 = \tfrac{\partial}{\partial R_{22}}\mathbb{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix};\\
&\mathbb{X}_9 = \tfrac{\partial}{\partial R_{33}}\mathbb{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\end{aligned} \tag{7.103}$$

Let us rewrite the nine independent matrix elements as u_i $(i = 1,2,\ldots,9)$ and the partial derivatives of $\mathbb{X}$ to u_i as $\mathbb{X}_i$:

$$\mathbb{X}_i = \frac{\partial}{\partial u_i}\mathbb{X}. \tag{7.104}$$

It then follows that

$$\mathbb{X} = \sum_{i=1}^{9} u_i \frac{\partial}{\partial u_i}\mathbb{X} = \sum_{i=1}^{9} u_i \mathbb{X}_i. \tag{7.105}$$

Eq. (7.88) then becomes

$$\mathbb{U} = \exp\left[i \sum_{i=1}^{9} u_i \mathbb{X}_i \right]. \tag{7.106}$$

All the elements of U(3) can thus be *generated* via this formula; that is, by choosing different sets of nine real values for the independent parameters u_i, we can form every possible unitary matrix $\mathbb{U}$ of U(3). In this way, every unitary matrix $\mathbb{U}$ is dependent on nine real parameters, and the unitary group U(3) is said to be a nine-parametric Lie group or a Lie group of dimension 9.

The nine independent Hermitian matrices $\mathbb{X}_i$ are the *generators* of U(3). These generators span the corresponding Lie algebra, denoted $\mathfrak{u}(3)$. The link between the elements of the U(3) group and the generators of the $\mathfrak{u}(3)$ algebra is given by the following *exponential map*.

Definition 7.4 (U(3) exponential map): There exists an *exponential map* $\exp : \mathfrak{u}(3) \rightarrow \mathrm{U}(3)$, given by

$$\mathbb{U} = \exp\left[i \sum_{i=1}^{9} u_i \mathbb{X}_i \right], \tag{7.107}$$

which relates the generators of the $\mathfrak{u}(3)$ *Lie algebra* with the elements of the U(3) *Lie group* by expressing the unitary matrix $\mathbb{U}$ in terms of the infinitesimal generators $\mathbb{X}_i$, as defined in Eq. (7.104). ■

The unitary group U(1)

As another simple example, let us consider the unitary group U(1), which consists of all unitary matrices $\mathbb{U}$ of dimension 1×1. A study of this group will prove useful in §7.4.4. Because of the unit dimensionality of $\mathbb{U}$, each matrix $\mathbb{U}$ can be written as a *number* U. In line with Eq. (7.88), we express these elements as

$$U = e^{i\mathbb{X}}, \tag{7.108}$$

with $\mathbb{X}$ a Hermitian matrix of dimension 1. That is, $\mathbb{X}$ is also a number. As a consequence of the Hermitian character of $\mathbb{X}$ (cf. Eqs. (7.93), (7.94), and (7.95)), $\mathbb{X}$ has to be a *real* number, which we denote by R:

$$\mathbb{X} = R. \tag{7.109}$$

The elements U of U(1) can then be written as

$$U = e^{iR}. \tag{7.110}$$

The corresponding $\mathfrak{u}(1)$ Lie algebra is spanned by $n^2 = 1^2 = 1$ generator. With the help of Eq. (7.104), this generator can be easily obtained from the general Hermitian matrix $\mathbb{X} = R$ with one independent matrix element R:

$$\frac{\partial}{\partial R} R = 1 = \mathbb{I}, \tag{7.111}$$

with $\mathbb{I}$ being the unit matrix of dimension 1. There thus exists an *exponential map* $\exp : \mathfrak{u}(1) \to \mathrm{U}(1)$, given by

$$U = \exp\left[iR\,\mathbb{I}\right] \tag{7.112}$$

that relates the infinitesimal generator $\mathbb{I}$ of the $\mathfrak{u}(1)$ *Lie algebra* with the elements U of the U(1) *Lie group*.

Finally, let us consider what happens when we act with one of the elements U from the group $U(1)$ on a wave function $|\psi\rangle$:

$$U\,|\psi\rangle = e^{iR}\,|\psi\rangle. \tag{7.113}$$

Clearly, this has the effect of multiplying the ket by an overall *phase factor* e^{iR}. Phase factors do not affect the diagonal matrix elements of a Hermitian operator $\hat{A}$. The expectation value for $\hat{A}$, for instance, remains invariant under the previous transformation (Eq. (7.113)), because the phase factors cancel out when we determine the measurable quantity $\langle \hat{A} \rangle$:

$$\langle\psi|\,\hat{A}\,|\psi\rangle = \langle\psi|\,U^{\dagger}\hat{A}U\,|\psi\rangle = \langle\psi|\,e^{-iR}\hat{A}e^{iR}\,|\psi\rangle. \tag{7.114}$$

We will come back to this point when discussing the relative importance of the groups U(3) and SU(3). For now, however, we first need to derive the generators of SU(3).

7.4.4 The generators of SU(3)

In comparison with U(3), the elements of the special unitary group SU(3) have one more restriction—namely, that their determinant must equal +1 (cf. Definition 7.2). The number of independent parameters, therefore, lowers from nine to eight, and SU(3) is said to be an eight-parametric Lie group. Its corresponding Lie algebra $\mathfrak{su}(3)$ is spanned by eight linearly independent generators, which we denote by the symbol $\mathbb{X}'_i$ $(i = 1, 2, \ldots, 8)$.

This raises a number of interesting questions. First of all, what is the influence of this restriction (unit determinant) on the general form of the SU(3) generators? Second, what is the relationship between the nine generators $\mathbb{X}_i$ of U(3) and the eight generators $\mathbb{X}'_i$ of SU(3)?

Traceless Hermitian matrices $\mathbb{X}'_i$

To formulate an answer to the first question, let us start by recalling the Jacobi formula for the determinant of a matrix exponential[16]:

$$\det\left(e^{\mathbb{C}}\right) = e^{\mathrm{Tr}(\mathbb{C})}, \tag{7.115}$$

with $\mathbb{C}$ representing any complex square matrix of dimension n, and $\mathrm{Tr}(\mathbb{C})$ denoting the *trace* of $\mathbb{C}$, as defined in Eq. (7.96). Using Eq. (7.88), and denoting the elements of SU(3) by the general symbol $\mathbb{U}'$, we write

$$\mathbb{U}' = e^{i\mathbb{X}'}. \tag{7.116}$$

Applying this to Eq. (7.115) yields

$$\det \mathbb{U}' = \det\left(e^{i\mathbb{X}'}\right) = e^{i\mathrm{Tr}(\mathbb{X}')}. \tag{7.117}$$

Since all elements of SU(3) have unit determinant ($\det \mathbb{U}' = +1$), it follows that

$$e^{i\mathrm{Tr}(\mathbb{X}')} = 1 \quad \rightarrow \quad \mathrm{Tr}\left(\mathbb{X}'\right) = 0. \tag{7.118}$$

We conclude from Eq. (7.104) that the eight generators $\mathbb{X}'_i$ of SU(3) are all *traceless* Hermitian matrices of order 3:

$$\mathrm{Tr}\left(\mathbb{X}'_i\right) = 0 \qquad \forall i = 1, 2, \ldots, 8. \tag{7.119}$$

Definition 7.5 (SU(3) exponential map): There exists an *exponential map* $\exp : \mathfrak{su}(3) \rightarrow \mathrm{SU}(3)$, given by

$$\mathbb{U}' = \exp\left[i \sum_{i=1}^{8} u'_i \mathbb{X}'_i \right], \tag{7.120}$$

that relates the generators of the $\mathfrak{su}(3)$ *Lie algebra* with the elements of the SU(3) *Lie group* by expressing the unitary matrix $\mathbb{U}'$ in terms of the infinitesimal generators $\mathbb{X}'_i$ with trace equal to zero. ▪

The generators of SU(3)

This brings us to the second question: How can we derive the eight traceless generators $\mathbb{X}'_i$ of SU(3) from the nine generators $\mathbb{X}_i$ of U(3)? This question leads us to the famous matrices of the *eightfold way* used by Gell-Mann in 1961 as a basis for the standard model of elementary particles (see §7.5). For this, we have to consider the eight generators of the SU(3) group that we will obtain from the nine U(3) generators by the method of trace reduction. The SU(3) generators must satisfy the property in Eq. (7.119); that is, all the generators $\mathbb{X}_i$ ($i = 1, 2, \ldots, 8$) must be *traceless* Hermitian matrices of order 3 (*vide supra*).

---

[16] The proof of this identity is in Appendix E.

Glancing at the matrices in Eq. (7.103), we observe that $\mathbb{X}_1$, $\mathbb{X}_2$, $\mathbb{X}_4$, $\mathbb{X}_5$, $\mathbb{X}_6$, and $\mathbb{X}_7$ already have a vanishing trace. Let us denote them by the new symbol λ_i:

$$\lambda_i \equiv \mathbb{X}_i \qquad i = 1,\ 2,\ 4,\ 5,\ 6, \text{ and } 7. \tag{7.121}$$

The remaining three matrices $\mathbb{X}_3$, $\mathbb{X}_8$, and $\mathbb{X}_9$—corresponding to parameters R_{11}, R_{22}, and R_{33}—do not satisfy Eq. (7.119) since they have $+1$ on the diagonal. Let us therefore carry out a transformation to a new set of two *traceless* Hermitian generators (denoted λ_3 and λ_8) and the 3×3 unit matrix $\mathbb{I}$ (denoted here as λ_9) by forming linearly independent combinations of the original generators $\mathbb{X}_3$, $\mathbb{X}_8$, and $\mathbb{X}_9$. Although there are numerous ways of recombining the original set $\{\mathbb{X}_3, \mathbb{X}_8, \mathbb{X}_9\}$ into a traceless set $\{\lambda_3, \lambda_8\}$ and the unit matrix $\{\lambda_9\}$, let us follow the standard method from elementary particle physics (as first introduced by Gell-Mann in 1961), and define λ_3, λ_8, and λ_9 as follows:

$$\begin{aligned}
\lambda_3 &\equiv \mathbb{X}_3 - \mathbb{X}_8 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \\
\lambda_8 &\equiv \frac{1}{\sqrt{3}} (\mathbb{X}_3 + \mathbb{X}_8 - 2\mathbb{X}_9) = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}; \\
\lambda_9 &\equiv \mathbb{X}_3 + \mathbb{X}_8 + \mathbb{X}_9 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\end{aligned} \tag{7.122}$$

This linear recombination effectively redistributes the nine generators of U(3) in a singleton (the unit matrix λ_9, playing the role of generator for the U(1) Lie group) and a set of eight traceless and Hermitian matrices λ_i ($i = 1 \to 8$), which form the generators of the special unitary Lie group SU(3). In the following, we will focus our attention on the generators of the SU(3) group only. These are the famous *Gell-Mann matrices* given in Table 7.5. They form the *fundamental* or *defining* representation of the SU(3) group. The $\sqrt{3}$ in the definition of λ_8 takes care that all eight matrices have the same weight (i.e., if we add the squared norms of all numbers in a given matrix, we always obtain two). We thus say that the generators λ_i obey the *normalization relation*

$$\operatorname{Tr}\left(\lambda_i \lambda_j\right) = 2\delta_{ij}, \tag{7.123}$$

which can be easily verified by explicit matrix calculation.

The factorization of U(3)

The nine Gell-Mann matrices, corresponding to the λ's in Table 7.5, together generate the group U(3). We can thus rewrite Eq. (7.106) as

$$\mathbb{U} = \exp\left[i \sum_{i=1}^{9} v_i \lambda_i \right], \tag{7.124}$$

where v_i denotes the independent parameter associated with the λ_i generator. This expression can now be *factorized* by splitting the sum over λ_1 through λ_8, which

Table 7.5 The Gell-Mann 3×3 matrices λ_i ($i = 1 \rightarrow 9$). These matrices are the generators of the unitary Lie group in three dimensions U(3). The first eight λ_i's form the generators for the *special* unitary Lie group SU(3), whereas the unit matrix λ_9 is the generator of U(1).

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$\lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}; \quad \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}; \quad \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix};$$

$$\lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}; \quad \lambda_8 = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{-2}{\sqrt{3}} \end{bmatrix}; \quad \lambda_9 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

generates SU(3) and the λ_9 singleton. Hence,

$$\mathbb{U} = \exp\left[i \sum_{i=1}^{8} v_i \lambda_i \right] \exp(i v_9 \lambda_9). \tag{7.125}$$

The first factor is the exponential map of SU(3); the second factor can be rewritten as

$$\exp(i v_9 \lambda_9) = \exp(i v_9 \mathbb{I}) = \mathbb{I} \exp(i v_9). \tag{7.126}$$

Putting these results together, we get

$$\mathbb{U} = \mathbb{U}' \, \mathbb{I} \exp(i v_9) = \mathbb{U}' \exp(i v_9). \tag{7.127}$$

Every element $\mathbb{U}$ of U(3) can thus be written as a product between $\mathbb{U}'$ (an element from the SU(3) group) and a phase factor (an element from the U(1) group). The full unitary group U(3) can thus be *factorized* as the *direct product*:

$$\mathrm{U}(3) = \mathrm{SU}(3) \otimes \mathrm{U}(1), \tag{7.128}$$

with $\otimes$ denoting the direct product.

7.4.5 The $\mathfrak{su}(3)$ Lie algebra

Having determined the eight generators λ_i of the SU(3) group, we can calculate their commutation relations as follows:

$$\left[\lambda_i, \lambda_j\right] = \lambda_i \lambda_j - \lambda_j \lambda_i. \tag{7.129}$$

After a number of tedious but elementary matrix computations and reductions, we obtain the results shown in Table 7.6. This table contains a wealth of information that we shall study in the next paragraphs. For a start, it is clear from Table 7.6 that the

Table 7.6 Commutation table for the generators λ_i of the $\mathfrak{su}(3)$ Lie algebra. The commutator $[\lambda_i, \lambda_j]$ of any λ_i and λ_j $(i, j = 1 \to 9)$ is listed in the ith row and jth column. Notice that for most commutators $[\lambda_i, \lambda_j]$, the result reduces to a single generator λ_k, whereas for the commutators $[\lambda_4, \lambda_5]$ and $[\lambda_6, \lambda_7]$, the result reduces to two summands—for example, $[\lambda_4, \lambda_5] = i\lambda_3 + \sqrt{3}i\lambda_8$. The commutation table is *skew-symmetric* as a consequence of the fundamental commutation relation $[A, B] = -[B, A]$. Thus, $\left[\lambda_i, \lambda_j\right] = -\left[\lambda_j, \lambda_i\right], \forall i, j = 1 \to 9$.

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9
λ_1	0	$2i\lambda_3$	$-2i\lambda_2$	$i\lambda_7$	$-i\lambda_6$	$i\lambda_5$	$-i\lambda_4$	0	0
λ_2	$-2i\lambda_3$	0	$2i\lambda_1$	$i\lambda_6$	$i\lambda_7$	$-i\lambda_4$	$-i\lambda_5$	0	0
λ_3	$2i\lambda_2$	$-2i\lambda_1$	0	$i\lambda_5$	$-i\lambda_4$	$-i\lambda_7$	$i\lambda_6$	0	0
λ_4	$-i\lambda_7$	$-i\lambda_6$	$-i\lambda_5$	0	$i\lambda_3 + \sqrt{3}i\lambda_8$	$i\lambda_2$	$i\lambda_1$	$-\sqrt{3}i\lambda_5$	0
λ_5	$i\lambda_6$	$-i\lambda_7$	$i\lambda_4$	$-i\lambda_3 - \sqrt{3}i\lambda_8$	0	$-i\lambda_1$	$i\lambda_2$	$\sqrt{3}i\lambda_4$	0
λ_6	$-i\lambda_5$	$i\lambda_4$	$i\lambda_7$	$-i\lambda_2$	$i\lambda_1$	0	$-i\lambda_3 + \sqrt{3}i\lambda_8$	$-\sqrt{3}i\lambda_7$	0
λ_7	$i\lambda_4$	$i\lambda_5$	$-i\lambda_6$	$-i\lambda_1$	$-i\lambda_2$	$i\lambda_3 - \sqrt{3}i\lambda_8$	0	$\sqrt{3}i\lambda_6$	0
λ_8	0	0	0	$\sqrt{3}i\lambda_5$	$-\sqrt{3}i\lambda_4$	$\sqrt{3}i\lambda_7$	$-\sqrt{3}i\lambda_6$	0	0
λ_9	0	0	0	0	0	0	0	0	0

Table 7.7 Table of nonvanishing structure constants f_{ijk} for the $\mathfrak{su}(3)$ Lie algebra. The f_{ijk} are totally *antisymmetric* under the exchange of any two indices. All nonvanishing structure constants can therefore be derived by permutation of the indices in the structure constants listed here.

ijk	f_{ijk}	ijk	f_{ijk}
123	1	147	$\frac{1}{2}$
156	$-\frac{1}{2}$	246	$\frac{1}{2}$
257	$\frac{1}{2}$	345	$\frac{1}{2}$
367	$-\frac{1}{2}$	458	$\frac{\sqrt{3}}{2}$
678	$\frac{\sqrt{3}}{2}$		

$\mathfrak{su}(3)$ Lie algebra is *closed* under commutation; that is, every commutator $[\lambda_i, \lambda_j]$ can be written as a linear combination of the eight generators:

$$\left[\lambda_i, \lambda_j\right] = \sum_{k=1}^{8} 2if_{ijk}\lambda_k \equiv 2if_{ijk}\lambda_k, \tag{7.130}$$

where we used Einstein's summation convention in the last equality. The constant factor $2i$ has been extracted for simplicity. The nonvanishing *structure constants* f_{ijk} are listed in Table 7.7 and are totally *antisymmetric* under the exchange of any two indices; that is,

$$f_{ijk} = -f_{jik} = f_{jki} = -f_{ikj} = \text{and so on.} \tag{7.131}$$

Definition 7.6 (The $\mathfrak{su}(3)$ Lie algebra): The $\mathfrak{su}(3)$ Lie algebra is the linear vector space, spanned by the generators λ_i ($i = 1 \rightarrow 8$), in which the product rule is defined as follows: $[\cdot,\cdot] : \mathfrak{su}(3) \times \mathfrak{su}(3) \rightarrow \mathfrak{su}(3)$. The multiplication operator $[\cdot,\cdot]$ is called the *Lie bracket*, for which $\left[\lambda_i, \lambda_j\right] = 2if_{ijk}\lambda_k$, with f_{ijk} the different *structure constants.* ■

Example: The commutator of λ_4 and λ_5

Let us, by way of example, compute the commutator $[\lambda_4, \lambda_5]$ of λ_4 and λ_5. The matrix form of these generators is given in Table 7.5. Since $[\lambda_4, \lambda_5] = \lambda_4\lambda_5 – \lambda_5\lambda_4$, we start by calculating the matrix products $\lambda_4\lambda_5$ and $\lambda_5\lambda_4$ respectively:

$$\begin{aligned} \lambda_4\lambda_5 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i \end{bmatrix}; \\ \lambda_5\lambda_4 &= \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & i \end{bmatrix}. \end{aligned} \tag{7.132}$$

The commutator $[\lambda_4,\lambda_5]$ is then found to equal

$$\begin{aligned}[\lambda_4,\lambda_5] &= \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & i \end{bmatrix} = \begin{bmatrix} 2i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2i \end{bmatrix} \\ &= i\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \sqrt{3}i\frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} = i\lambda_3 + \sqrt{3}i\lambda_8,\end{aligned} \tag{7.133}$$

as summarized in Table 7.6. Rewriting this as

$$[\lambda_4,\lambda_5] = 2i\left[\frac{1}{2}\lambda_3 + \frac{\sqrt{3}}{2}\lambda_8\right], \tag{7.134}$$

we can read off the following structure constants:

$$\begin{aligned} f_{453} &= \frac{1}{2}; \\ f_{458} &= \frac{\sqrt{3}}{2}, \end{aligned} \tag{7.135}$$

which is in agreement with the values for the structure constants shown in Table 7.7.[17]

7.4.6 The Cartan subalgebra of $\mathfrak{su}(3)$

With all this in hand, we are ready to get down to the serious nitty-gritty of the $\mathfrak{su}(3)$ algebra. Although the structure constants f_{ijk} were seen to define the $\mathfrak{su}(3)$ algebra completely, let us delve even deeper into the variegated anatomy of the $\mathfrak{su}(3)$ algebra to reveal its inner workings. One way of uncovering the full power of the $\mathfrak{su}(3)$ algebra is by introducing its *Cartan subalgebra* $\mathfrak{H}$ and associated *Weyl diagram*. To this aim, we have to change the basis of the $\mathfrak{su}(3)$ algebra, given by the Gell-Mann matrices in Table 7.5, and derive the more useful *Cartan-Weyl basis*. For this, we shall proceed along the threefold path as outlined at the end of Chapter 6, §6.4.

Step 1: The Cartan subalgebra and Cartan generators

To start with, let us look for the maximal subset of mutually commuting generators $\lambda_i \in \mathfrak{su}(3)$. On closer inspection of Table 7.6, we find the following three commutators:

$$[\lambda_1,\lambda_8] = 0, \qquad [\lambda_2,\lambda_8] = 0, \text{ and } \qquad [\lambda_3,\lambda_8] = 0. \tag{7.136}$$

Since λ_1, λ_2, and λ_3 do *not* commute among themselves,

$$[\lambda_1,\lambda_2] \neq 0, \qquad [\lambda_1,\lambda_3] \neq 0, \text{ and } \qquad [\lambda_2,\lambda_3] \neq 0, \tag{7.137}$$

not more than two λ's can be diagonalized simultaneously. From the possible pairs $\{\lambda_1,\lambda_8\}$, $\{\lambda_2,\lambda_8\}$, and $\{\lambda_3,\lambda_8\}$, let us chose the last pair, $\{\lambda_3,\lambda_8\}$, as our commuting set. It is convenient to rewrite both operators in a slightly different form:

$$\hat{T}_3 \equiv \frac{1}{2}\lambda_3; \qquad \hat{Y} \equiv \frac{1}{\sqrt{3}}\lambda_8. \tag{7.138}$$

---

[17] Note that f_{453} can be obtained from the structure constant $f_{345} = 1/2$ listed in Table 7.7 via two permutations of the indices, corresponding to two consecutive changes of sign: $f_{345} = 1/2 \rightarrow f_{435} = -1/2 \rightarrow f_{453} = 1/2$.

Both operators will reappear later as the isospin and hypercharge operators, respectively, in our examination of the eightfold way. Since λ_3 and λ_8 have simply been *rescaled*, their commuting property has not been upset and

$$\left[\hat{T}_3, \hat{Y}\right] = 0. \tag{7.139}$$

The set $\{\hat{T}_3, \hat{Y}\}$ thus forms a basis for a *maximal Abelian subalgebra* $\mathfrak{H} \subset \mathfrak{su}(3)$. This is called the *Cartan subalgebra* of $\mathfrak{su}(3)$. The operators $\hat{T}_3$ and $\hat{Y}$ are referred to as *Cartan generators*, and the dimension 2 of $\mathfrak{H}$ defines the *rank* of the $\mathfrak{su}(3)$ Lie algebra.[18]

In view of their commuting character, $\hat{T}_3$ and $\hat{Y}$ are simultaneously *diagonalizable*. We denote their eigenvalues by T_3 and Y, respectively, and represent their common eigenstates by the ket $|T_3, Y\rangle$. This then yields the following eigenvalue equations:

$$\hat{T}_3 \,|T_3, Y\rangle = T_3 \,|T_3, Y\rangle; \tag{7.140}$$

$$\hat{Y} \,|T_3, Y\rangle = Y \,|T_3, Y\rangle. \tag{7.141}$$

The T_3 and Y eigenvalues are referred to as the *weights* of the Cartan generators $\hat{T}_3$ and $\hat{Y}$, and are used to label the substates within an SU(3) multiplet. In a *Weyl diagram*, the T_3 and Y eigenvalues function as coordinates to plot the substates of a given multiplet in the T_3—Y-plane (*vide infra*). The weights T_3 and Y can then be considered as the components of a two-dimensional *weight vector* **h**, which sets in at the origin O and points to its corresponding substate.

Step 2: Weyl generators and Weyl diagrams

In a second step, we arrange the remaining generators λ_i of $\mathfrak{su}(3)$ $(i = 1,2,4,5,6,7)$ into linear combinations to form a linearly independent set of *raising* and *lowering* operators:

$$\begin{aligned}
\hat{T}_+ &\equiv \frac{1}{2}(\lambda_1 + i\lambda_2); & \hat{T}_- &\equiv \frac{1}{2}(\lambda_1 - i\lambda_2);\\
\hat{U}_+ &\equiv \frac{1}{2}(\lambda_6 + i\lambda_7); & \hat{U}_- &\equiv \frac{1}{2}(\lambda_6 - i\lambda_7);\\
\hat{V}_+ &\equiv \frac{1}{2}(\lambda_4 + i\lambda_5); & \hat{V}_- &\equiv \frac{1}{2}(\lambda_4 - i\lambda_5).
\end{aligned} \tag{7.142}$$

In the above definition, we have clearly been inspired by the general form of the angular momentum ladder operators $\hat{L}_\pm \equiv \hat{L}_x \pm i\hat{L}_y$, as defined in Chapter 5. The six ladder operators in Eq. (7.142) are referred to as *Weyl generators*. Along with the two Cartan generators $\hat{T}_3$ and $\hat{Y}$, they constitute the *Cartan-Weyl basis* for the $\mathfrak{su}(3)$ algebra:

$$\{\hat{T}_3,\ \hat{Y},\ \hat{T}_+,\ \hat{T}_-,\ \hat{U}_+,\ \hat{U}_-,\ \hat{V}_+,\ \hat{V}_-\}. \tag{7.143}$$

If we denote the Cartan generators for a moment by the customary symbol $\hat{H}_i$ $(i = 1,2)$ and the Weyl generators by the symbol $\hat{E}_\alpha$ $(\alpha = 1,\ldots,6)$, the Cartan-Weyl basis can be written more generally as

$$\{\hat{H}_1,\ \hat{H}_2,\ \hat{E}_1,\ \hat{E}_2,\ \hat{E}_3,\ \hat{E}_4,\ \hat{E}_5,\ \hat{E}_6\}. \tag{7.144}$$

---

[18] It can be shown more generally that an $\mathfrak{su}(n)$ algebra is a Lie algebra of rank $(n-1)$.

This shows more clearly that the $\mathfrak{su}(3)$ algebra has actually been decomposed into a direct sum of the Cartan subalgebra $\mathfrak{H}$ (spanned by $\hat{H}_1$ and $\hat{H}_2$) and six one-dimensional subalgebras $\mathfrak{E}_\alpha$, generated by the Weyl generators $\hat{E}_\alpha$:

$$\mathfrak{su}(3) = \mathfrak{H} \bigoplus_{\alpha=1}^{6} \mathfrak{E}_\alpha = \mathfrak{H} \oplus \mathfrak{E}_1 \oplus \ldots \oplus \mathfrak{E}_5 \oplus \mathfrak{E}_6. \tag{7.145}$$

Returning to our discussion of the Weyl generators, and the fact that they act as *step operators*, they must satisfy the general commutation relations:

$$\left[\hat{H}_i, \hat{E}_\alpha\right] = \alpha_i \hat{E}_\alpha, \qquad \forall i = 1,2;\ \alpha = 1,\ldots,6. \tag{7.146}$$

That is to say, the Weyl generators $\hat{E}_\alpha$ must behave as *eigenoperators* of the Cartan generators $\hat{H}_i$. Recall that the eigenvalues α_i are referred to as the *roots* of $\hat{E}_\alpha$ with respect to $\hat{H}_i$.

Let us verify the validity of Eq. (7.146) by explicitly calculating the different commutators. With the help of the commutation table, Table 7.6, and the defining expressions for the Cartan and Weyl generators in Eqs. (7.138) and (7.142), the commutator of $\hat{T}_3$ and $\hat{T}_+$ can be written as

$$\begin{aligned}
\left[\hat{T}_3,\ \hat{T}_+\right] &= \left[\frac{1}{2}\lambda_3,\ \frac{1}{2}(\lambda_1 + i\lambda_2)\right] \\
&= \frac{1}{4}\left([\lambda_3,\lambda_1] + i[\lambda_3,\lambda_2]\right) \\
&= \frac{1}{4}(2i\lambda_2 + 2\lambda_1) \\
&= \frac{1}{2}(\lambda_1 + i\lambda_2) \\
\left[\hat{T}_3,\ \hat{T}_+\right] &= \hat{T}_+.
\end{aligned} \tag{7.147}$$

This establishes $\hat{T}_+$ as an eigenoperator of $\hat{T}_3$ with root $+1$. The other commutators can be worked out in a similar vein. Here, we only mention the results:

$$\begin{aligned}
\left[\hat{T}_3,\ \hat{T}_+\right] &= +\hat{T}_+; & \left[\hat{Y},\ \hat{T}_+\right] &= 0; \\
\left[\hat{T}_3,\ \hat{T}_-\right] &= -\hat{T}_-; & \left[\hat{Y},\ \hat{T}_-\right] &= 0; \\
\left[\hat{T}_3,\ \hat{U}_+\right] &= -\tfrac{1}{2}\hat{U}_+; & \left[\hat{Y},\ \hat{U}_+\right] &= +\hat{U}_+; \\
\left[\hat{T}_3,\ \hat{U}_-\right] &= +\tfrac{1}{2}\hat{U}_-; & \left[\hat{Y},\ \hat{U}_-\right] &= -\hat{U}_-; \\
\left[\hat{T}_3,\ \hat{V}_+\right] &= +\tfrac{1}{2}\hat{V}_+; & \left[\hat{Y},\ \hat{V}_+\right] &= +\hat{V}_+; \\
\left[\hat{T}_3,\ \hat{V}_-\right] &= -\tfrac{1}{2}\hat{V}_-; & \left[\hat{Y},\ \hat{V}_-\right] &= -\hat{V}_-.
\end{aligned} \tag{7.148}$$

(The commutation relations among the Weyl generators themselves are listed in Table 7.8.) According to Eq. (7.148), every Weyl generator acts as an eigenoperator of $\hat{T}_3$ and $\hat{Y}$.

The importance of this statement is the following: when one of the Weyl generators acts on the ket $|T_3, Y\rangle$, it shifts the eigenvalues T_3 and Y by an amount given by the roots of that Weyl generator with respect to the Cartan generators $\hat{T}_3$ and $\hat{Y}$. To see

Table 7.8 Commutation table for the generators of the $\mathfrak{su}(3)$ Lie algebra in the *Cartan-Weyl basis*. The commutator $[X_i, X_j]$ $(i, j = 1 \rightarrow 8)$ is listed in the ith row and jth column. The commutation table is *skew-symmetric* as a consequence of the fundamental commutation relation $\left[X_i, X_j\right] = -\left[X_j, X_i\right]$. The *Cartan subalgebra* $\mathfrak{H} = \{\hat{Y}, \hat{T}_3\}$ forms a maximal Abelian subalgebra of $\mathfrak{su}(3)$, as indicated by the square of bold zeros in the upper left corner of the table.

	$\hat{Y}$	$\hat{T}_3$	$\hat{T}_+$	$\hat{T}_-$	$\hat{U}_+$	$\hat{U}_-$	$\hat{V}_+$	$\hat{V}_-$
$\hat{Y}$	$\mathbf{0}$	$\mathbf{0}$	0	0	$\hat{U}_+$	$-\hat{U}_-$	$\hat{V}_+$	$-\hat{V}_-$
$\hat{T}_3$	$\mathbf{0}$	$\mathbf{0}$	$\hat{T}_+$	$-\hat{T}_-$	$-\frac{1}{2}\hat{U}_+$	$\frac{1}{2}\hat{U}_-$	$\frac{1}{2}\hat{V}_+$	$-\frac{1}{2}\hat{V}_-$
$\hat{T}_+$	0	$-\hat{T}_+$	0	$2\hat{T}_3$	$\hat{V}_+$	0	0	$-\hat{U}_-$
$\hat{T}_-$	0	$\hat{T}_-$	$-2\hat{T}_3$	0	0	$-\hat{V}_-$	$\hat{U}_+$	0
$\hat{U}_+$	$-\hat{U}_+$	$\frac{1}{2}\hat{U}_+$	$-\hat{V}_+$	0	0	$\frac{3}{2}\hat{Y} - \hat{T}_3$	0	$\hat{T}_-$
$\hat{U}_-$	$\hat{U}_-$	$-\frac{1}{2}\hat{U}_-$	0	$\hat{V}_-$	$-\frac{3}{2}\hat{Y} + \hat{T}_3$	0	$-\hat{T}_+$	0
$\hat{V}_+$	$-\hat{V}_+$	$-\frac{1}{2}\hat{V}_+$	0	$-\hat{U}_+$	0	$\hat{T}_+$	0	$\frac{3}{2}\hat{Y} + \hat{T}_3$
$\hat{V}_-$	$\hat{V}_-$	$\frac{1}{2}\hat{V}_-$	$\hat{U}_-$	0	$-\hat{T}_-$	0	$-\frac{3}{2}\hat{Y} - \hat{T}_3$	0

how this works, consider the action of the raising operator $\hat{T}_+$ on the eigenket $|T_3, Y\rangle$ using Eq. (7.148):

$$\begin{aligned} \hat{T}_3 \hat{T}_+ |T_3, Y\rangle &= \left(\left[\hat{T}_3, \hat{T}_+\right] + \hat{T}_+ \hat{T}_3\right) |T_3, Y\rangle \\ &= \left(\hat{T}_+ + \hat{T}_+ T_3\right) |T_3, Y\rangle \\ &= (T_3 + 1)\, \hat{T}_+ |T_3, Y\rangle. \end{aligned} \tag{7.149}$$

$\hat{T}_+$ raises the eigenvalue T_3 by an amount $+1$, which is the root of $\hat{T}_+$ with respect to $\hat{T}_3$ according to the commutation relations in Eq. (7.148). Similarly,

$$\begin{aligned} \hat{Y} \hat{T}_+ |T_3, Y\rangle &= \left(\left[\hat{Y}, \hat{T}_+\right] + \hat{T}_+ \hat{Y}\right) |T_3, Y\rangle \\ &= \left(0 + \hat{T}_+ Y\right) |T_3, Y\rangle \\ &= (Y + 0)\, \hat{T}_+ |T_3, Y\rangle. \end{aligned} \tag{7.150}$$

In this case, $\hat{T}_+$ leaves the eigenvalue Y of $\hat{Y}$ unchanged. The previous two equations can then be combined to yield

$$\hat{T}_+ |T_3, Y\rangle \rightarrow |T_3 + 1, Y\rangle. \tag{7.151}$$

By the same argument, we can deduce the action of every Weyl generator on the ket $|T_3, Y\rangle$:

$$\hat{T}_\pm |T_3, Y\rangle \rightarrow |T_3 \pm 1, Y\rangle; \tag{7.152}$$

$$\hat{U}_\pm |T_3, Y\rangle \rightarrow |T_3 \mp 1/2,\ Y \pm 1\rangle; \tag{7.153}$$

$$\hat{V}_\pm |T_3, Y\rangle \rightarrow |T_3 \pm 1/2,\ Y \pm 1\rangle. \tag{7.154}$$

We can also depict the actions of the Weyl generators in a *Weyl diagram*. For this, we consider the *roots* of every Weyl generator as the components of a *root vector* that lies

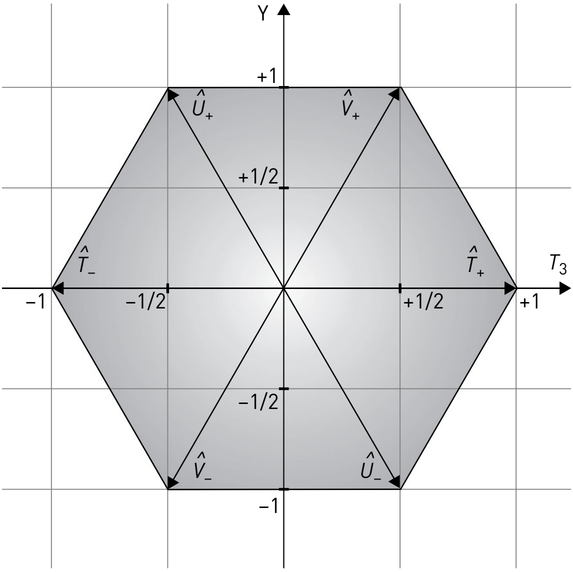


FIGURE 7.5 Root diagram of the $\mathfrak{su}(3)$ Lie algebra. The action of every Weyl generator is shown in the T_3—Y-plane. The Cartan generators $\hat{T}_3$ and $\hat{Y}$ are positioned at the origin of the Weyl diagram.

in a two-dimensional *weight space*. For example, the roots of $\hat{T}_+$ are $+1$ and 0 with respect to $\hat{T}_3$ and $\hat{Y}$, respectively. These form the components of a root vector, which we denote using the same operator symbol $\hat{T}_+$ for simplicity. After positioning all the root vectors in the T_3—Y-plane, we obtain the two-dimensional *root diagram* of the $\mathfrak{su}(3)$ Lie algebra, as shown in Figure 7.5. This is also called a *Weyl diagram*. Note that the end points of the Weyl generators correspond to the corners of a *regular hexagon*.[19] Besides the Weyl generators, we can also position the Cartan generators in the Weyl diagram. However, in view of their commuting property (Eq. (7.139)), the roots of the Cartan generators $\hat{T}_3$ and $\hat{Y}$ are all zero. The root vectors of the Cartan generators thus correspond to two *null vectors*.

Step 3: Casimir invariants

The $\mathfrak{su}(3)$ algebra is a Lie algebra of rank 2. According to Racah's theorem (Chapter 6, §6.4), this implies the existence of two independent *Casimir operators* $\hat{C}_\mu$ that commute with all the generators λ_i ($i = 1 \to 8$) of $\mathfrak{su}(3)$:

$$\left[\hat{C}_\mu, \lambda_i\right] = 0, \qquad \forall \mu = 1 \to 2,\ i = 1 \to 8. \tag{7.155}$$

The Casimir invariants $\hat{C}_1$ and $\hat{C}_2$ can be used to label the *irreducible representations* (i.e., multiplets) of the SU(3) group and are commonly defined as

$$\begin{aligned} \hat{C}_1 &= \sum_{i=1}^{8} \frac{1}{2}\lambda_i^2; \\ \hat{C}_2 &= \sum_{ijk} \frac{1}{8} d_{ijk}\lambda_i\lambda_j\lambda_k, \end{aligned} \tag{7.156}$$

---

[19] The units of the T_3- and Y-axes in the Weyl diagram have been scaled so as to form angles of 60° between the root vectors.

where d_{ijk} are constants originating from the *anticommutation* relations $\{\lambda_i, \lambda_j\} = \lambda_i\lambda_j + \lambda_j\lambda_i$ $(i, j = 1 \rightarrow 8)$.

7.4.7 The $\mathfrak{T}$, $\mathfrak{U}$, and $\mathfrak{V}$ subalgebras of $\mathfrak{su}(3)$

The $\mathfrak{T}$ subalgebra

In this section, we study a number of subalgebras of the $\mathfrak{su}(3)$ Lie algebra. We can easily establish from the commutation relations in Table 7.8 that operators $\hat{T}_3$, $\hat{T}_+$, and $\hat{T}_-$ form a closed subalgebra of $\mathfrak{su}(3)$:

$$\left[\hat{T}_3, \hat{T}_\pm\right] = \pm\hat{T}_\pm, \qquad \left[\hat{T}_+, \hat{T}_-\right] = 2\hat{T}_3. \tag{7.157}$$

Let us call this the $\mathfrak{T}$ *subalgebra*. Since the commutation relations among the $\hat{T}$ operators match those of the angular momentum algebra

$$\left[\hat{L}_3, \hat{L}_\pm\right] = \pm\hat{L}_\pm, \qquad \left[\hat{L}_+, \hat{L}_-\right] = 2\hat{L}_3, \tag{7.158}$$

we can identify the $\mathfrak{T}$ subalgebra as an $\mathfrak{so}(3)$ subalgebra of $\mathfrak{su}(3)$. This provides a connection between the unitary and orthogonal groups. That is, when reducing the symmetry group to rotational symmetry, the following *symmetry breaking* occurs:

$$\mathrm{SU}(3) \rightarrow \mathrm{SO}(3). \tag{7.159}$$

This corresponds to a transition from the $\mathfrak{su}(3)$ Lie algebra to the $\mathfrak{T}$ subalgebra that is isomorphic to $\mathfrak{so}(3)$. As a consequence, the rank lowers from 2 to 1 and $\hat{T}_3$ remains the only Cartan generator. This appears on the Cartan-Weyl diagram as a *projection* on the horizontal axis, which corresponds to the $\hat{T}_3$ operator (Figure 7.6). The octuplet of the SU(3) generators then gives rise to the spherical T_3 eigenvalues ±1, $\pm1/2$, $\pm1/2$, 0, 0. Clearly, this set corresponds to the signature of four spherical shells with $T = 0$, $T = 1$ and twice $T = 1/2$.

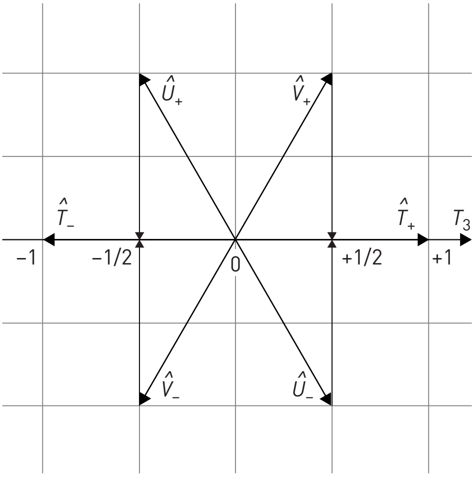


FIGURE 7.6 Root diagram showing the reduction of the $\mathfrak{su}(3)$ Lie algebra to the $\mathfrak{T}$ subalgebra.

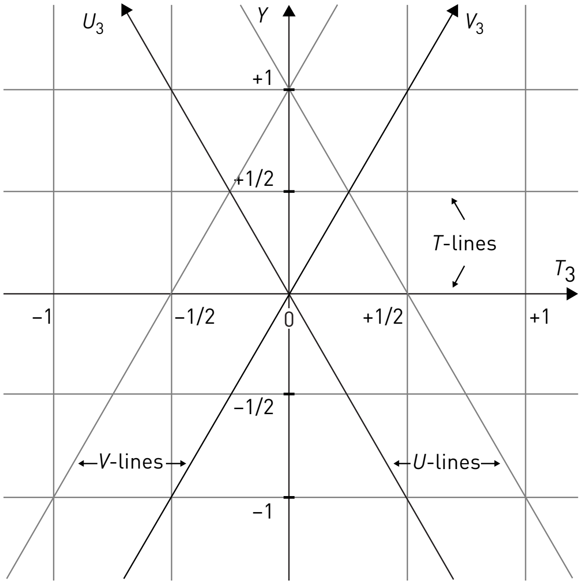


FIGURE 7.7 Root diagram showing the $\mathfrak{T}$, $\mathfrak{U}$, and $\mathfrak{V}$ subalgebras. The $\mathfrak{T}$, $\mathfrak{U}$, and $\mathfrak{V}$ submultiplets are located along the T-, U- and V-lines, respectively.

The $\mathfrak{U}$ and $\mathfrak{V}$ subalgebras

The sets of operators $\{\hat{U}_+, \hat{U}_-\}$ and $\{\hat{V}_+, \hat{V}_-\}$ do *not* form closed subalgebras, as can be seen from the commutation relations in Table 7.8:

$$\left[\hat{U}_+, \hat{U}_-\right] = \frac{3}{2}\hat{Y} - \hat{T}_3; \quad \left[\hat{V}_+, \hat{V}_-\right] = \frac{3}{2}\hat{Y} + \hat{T}_3. \tag{7.160}$$

However, in pursuing the analogy with the $\mathfrak{T}$ subalgebra, we might define two new diagonal operators $\hat{U}_3$ and $\hat{V}_3$:

$$2\hat{U}_3 \equiv \frac{3}{2}\hat{Y} - \hat{T}_3; \quad 2\hat{V}_3 \equiv \frac{3}{2}\hat{Y} + \hat{T}_3, \tag{7.161}$$

which yields a number of new commutation relations, as listed in Table 7.9. As a result, operators $\{\hat{U}_3, \hat{U}_+, \hat{U}_-\}$ and $\{\hat{V}_3, \hat{V}_+, \hat{V}_-\}$ are seen to form two additional $\mathfrak{so}(3)$ subalgebras: the $\mathfrak{U}$ and $\mathfrak{V}$ subalgebras, respectively (see Figure 7.7).

As is clear from the bold zeros in Table 7.8, the original set of mutually commuting generators $\{\hat{Y}, \hat{T}_3\}$ has now been enlarged to include the new operators $\hat{U}_3$ and $\hat{V}_3$: $\{\hat{Y}, \hat{T}_3, \hat{U}_3, \hat{V}_3\}$. However, this does *not* increase the rank of the $\mathfrak{su}(3)$ algebra from 2 to 4 because the $\hat{Y}$, $\hat{T}_3$ $\hat{U}_3$, and $\hat{V}_3$ operators are not all linearly independent. For instance,

$$\hat{U}_3 = -\hat{T}_3 + \hat{V}_3. \tag{7.162}$$

7.5 THE EIGHTFOLD WAY (1961)

Let us go back to the situation in particle physics at the beginning of the 1960s. The ever-growing particle zoo of elementary particles was causing much confusion and discussion among particle physicists. Little did they realize that with Gell-Mann's introduction of *strangeness* as a new conserved additive quantum number (see §7.1.5), the path was finally opened to a rational classification of the elementary particles.

Table 7.9 Extended commutation table for the generators of the $\mathfrak{su}(3)$ Lie algebra in the *Cartan-Weyl basis*. The commutator $[X_i, X_j]$ is listed in the ith row and jth column. The $\mathfrak{T}$, $\mathfrak{U}$, and $\mathfrak{V}$ *subalgebras* of $\mathfrak{su}(3)$ are located in the three square regions. Operators $\hat{Y}$, $\hat{T}_3$, $\hat{U}_3$, and $\hat{V}_3$ form a maximal Abelian subalgebra of $\mathfrak{su}(3)$, as indicated by the bold zeros in the table

	$\hat{Y}$	$\hat{T}_3$	$\hat{T}_+$	$\hat{T}_-$	$\hat{U}_3$	$\hat{U}_+$	$\hat{U}_-$	$\hat{V}_3$	$\hat{V}_+$	$\hat{V}_-$
$\hat{Y}$	$\mathbf{0}$	$\mathbf{0}$	0	0	$\mathbf{0}$	$\hat{U}_+$	$-\hat{U}_-$	$\mathbf{0}$	$\hat{V}_+$	$-\hat{V}_-$
$\hat{T}_3$	$\mathbf{0}$	$\mathbf{0}$	$\hat{T}_+$	$-\hat{T}_-$	$\mathbf{0}$	$-\frac{1}{2}\hat{U}_+$	$+\frac{1}{2}\hat{U}_-$	$\mathbf{0}$	$\frac{1}{2}\hat{V}_+$	$-\frac{1}{2}\hat{V}_-$
$\hat{T}_+$	0	$-\hat{T}_+$	0	$2\hat{T}_3$	$\frac{1}{2}\hat{T}_+$	$\hat{V}_+$	0	$-\frac{1}{2}\hat{T}_+$	0	$-\hat{U}_-$
$\hat{T}_-$	0	$\hat{T}_-$	$-2\hat{T}_3$	0	$-\frac{1}{2}\hat{T}_-$	0	$-\hat{V}_-$	$\frac{1}{2}\hat{T}_-$	$\hat{U}_+$	0
$\hat{U}_3$	$\mathbf{0}$	$\mathbf{0}$	$-\frac{1}{2}\hat{T}_+$	$\frac{1}{2}\hat{T}_-$	$\mathbf{0}$	$\hat{U}_+$	$-\hat{U}_-$	$\mathbf{0}$	$\frac{1}{2}\hat{V}_+$	$-\frac{1}{2}\hat{V}_-$
$\hat{U}_+$	$-\hat{U}_+$	$\frac{1}{2}\hat{U}_+$	$-\hat{V}_+$	0	$-\hat{U}_+$	0	$2\hat{U}_3$	$-\frac{1}{2}\hat{U}_+$	0	$\hat{T}_-$
$\hat{U}_-$	$\hat{U}_-$	$-\frac{1}{2}\hat{U}_-$	0	$\hat{V}_-$	$\hat{U}_-$	$-2\hat{U}_3$	0	$\frac{1}{2}\hat{U}_-$	$-\hat{T}_+$	0
$\hat{V}_3$	$\mathbf{0}$	$\mathbf{0}$	$\frac{1}{2}\hat{T}_+$	$-\frac{1}{2}\hat{T}_-$	$\mathbf{0}$	$\frac{1}{2}\hat{U}_+$	$-\frac{1}{2}\hat{U}_-$	$\mathbf{0}$	$\hat{V}_+$	$-\hat{V}_-$
$\hat{V}_+$	$-\hat{V}_+$	$-\frac{1}{2}\hat{V}_+$	0	$-\hat{U}_+$	$-\frac{1}{2}\hat{V}_+$	0	$\hat{T}_+$	$-\hat{V}_+$	0	$2\hat{V}_3$
$\hat{V}_-$	$\hat{V}_-$	$\frac{1}{2}\hat{V}_-$	$\hat{U}_-$	0	$\frac{1}{2}\hat{V}_-$	$-\hat{T}_-$	0	$\hat{V}_-$	$-2\hat{V}_3$	0

In this section, the enthralling story is told of how American physicist Murray Gell-Mann brought order to the jungle of elementary particles in 1961.[20] However, before venturing into a detailed account of the process by which Gell-Mann arrived at his classification schemes, we first have to introduce two more quantum numbers: the *hypercharge* Y and *baryon number* B. Both are closely related to strangeness, with hypercharge being defined as the sum of strangeness and baryon number:

$$Y \equiv S + B. \tag{7.163}$$

The baryon number takes on the values $+1$ for baryons, -1 for antibaryons, and 0 for all other particles (i.e., mesons and leptons).[21] Just like strangeness, both hypercharge and baryon number are strictly conserved quantum numbers in strong processes. Their values are listed in the Tables 7.1 through 7.4, along with the strangeness quantum number S, charge number Q, and isospin component T_3. A study of these values reveals that the charge number Q is given by

$$Q = T_3 + \frac{1}{2} Y = T_3 + \frac{1}{2} (S + B) \,. \tag{7.164}$$

This equation is known as the *Gell-Mann–Nishijima formula.* With this, all the ingredients are finally in place for our account of Gell-Mann's groundbreaking work.

7.5.1 An octet of particles

The pseudoscalar meson octet

The trick toward achieving a coherent systematization was to order all the known particles by their isospin component T_3 and hypercharge Y in a two-dimensional scheme. This process is illustrated in Figure 7.8 for the pseudoscalar mesons of Table 7.1 (except η'^0). The result, remarkably, was an *octet of particles*, with six particles lying at the corners of a regular hexagon and two more particles situated at the center of the graph. Inspired by this representation, Gell-Mann decided to name his scheme the *eightfold way*, in which he amusingly alluded to the *Noble Eightfold Path* of Buddhism toward enlightenment.

For each particle in Figure 7.8, the values of the hypercharge and isospin component can be read off the Y- and T_3-axes. Particles lying on a *horizontal line*, parallel to the T_3-axis, share the same value for strangeness; the three varieties of pions (π^-, π^0, and π^+), for instance, are characterized by strangeness $S = 0$. The kaons on the upper line (K^0 and K^+) have $S = +1$, whereas those on the lower line (K^- and $\bar{K}^0$) have $S = -1$.

In a similar vein, downward-sloping *diagonal lines* associate particles of like charge number Q; the neutral particles K^0, $\bar{K}^0$, π^0, and η^0, for example, are all situated along the main diagonal of the graph. Similarly, $Q = -1$ for π^- and K^-, and $Q = +1$ for K^+ and π^+. Particles taking diametrically opposite positions in the scheme are each other's *antiparticle.* We thus find the familiar particle–antiparticle pairs K^0–$\bar{K}^0$, K^+–K^-, and π^+–π^-. (Note that the particles π^0 and η^0 are their own antiparticles.)

---

[20] Israeli theoretical physicist Yuval Ne'eman (1925–2006) independently discovered a way of classifying the hadrons in 1961 based on the SU(3) flavor symmetry (*vide infra*).

[21] We return to the origin of these assignments in §7.6.2 and §7.6.3.

Interparticle transformations

You no doubt have observed the similarity between the hexagonal arrangement of pseudoscalar mesons in Figure 7.8 and the Weyl diagram for the $\mathfrak{su}(3)$ algebra shown in Figure 7.5. Gell-Mann realized this was not a lucky coincidence, nor a mere analogy, but a crucially important sign that the elementary particles were somehow governed by the unitary symmetry group SU(3). He proclaimed that the pseudoscalar mesons were grouped into a *supermultiplet* of the SU(3) group; the octet of particles was therefore considered to be a *representation* of the SU(3) group. Furthermore, Gell-Mann believed the strong force was *invariant* under the transformations of the SU(3) group. As a consequence, the strong nuclear force would not *distinguish* between the pseudoscalar mesons of the same octet and would therefore affect them the same way; the pseudoscalar mesons are said to be *symmetric* under the strong force. That is to say, the interchange of one meson for another goes *unnoticed* in strong interactions. Mathematically, this replacement can be described by the elements of the SU(3) group; the mesons can be transformed into each other under the action of the Weyl generators.

Let us clarify all this with a simple example. Suppose we would like to transform the neutral kaon K^0 into the eta meson η^0. As seen in Figure 7.8, the kaon K^0 has $T_3 = -1/2$ and $Y = +1$. We can therefore represent this particle by the ket

$$\left|K^0\right\rangle = |T_3, Y\rangle = \left|-\frac{1}{2}, +1\right\rangle. \tag{7.165}$$

Similarly, the eta meson η^0 can be represented by the ket

$$\left|\eta^0\right\rangle = |T_3, Y\rangle = |0, 0\rangle\,. \tag{7.166}$$

It then becomes evident that to transform K^0 into η^0, we need to simultaneously increase the T_3 quantum number by 1/2 (from $-1/2$ to 0) and lower the Y quantum number by 1 (from $+1$ to 0). From a consideration of the Weyl diagram in Figure 7.5, and Eqs. (7.152) through (7.154), it is clear this can be accomplished under the action of the $\hat{U}_-$ operator:

$$\hat{U}_- \left|K^0\right\rangle \rightarrow \left|\eta^0\right\rangle. \tag{7.167}$$

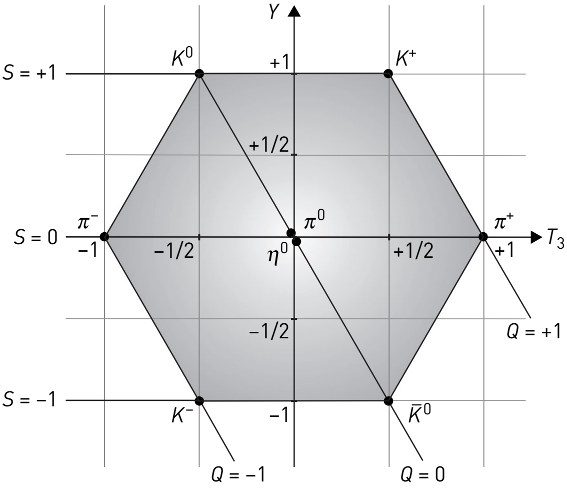


FIGURE 7.8 Weight diagram of the *pseudoscalar meson octet.*

Isospin multiplets

The SU(3) group contains an SO(3) subgroup, as explained in §7.4.7. When the SU(3) symmetry is broken to the spherical symmetry group SO(3), the SU(3) multiplets are split into a number of SO(3) submultiplets. These multiplets are characterized by two quantum numbers: (1) the total isospin T labeling the multiplet and (2) the isospin component T_3, distinguishing the various particle states of the multiplet. The SO(3) multiplets are called *isospin multiplets* whereas the SU(3) multiplets are referred to as *supermultiplets*. Particles on the same horizontal line in Figure 7.8 fall into such multiplets; each isospin multiplet therefore carries the same hypercharge Y and strangeness charge S.

The kaons K^0 and K^+, for instance, form an *isospin doublet* with $Y = +1$ and $S = +1$. Since the isospin component $T_3 = -1/2$ for K^0 and $T_3 = +1/2$ for K^+, the doublet is assigned a total isospin $T = 1/2$. The three pion varieties, on the other hand (π^-, π^0, and π^+), form an *isospin triplet* with $Y = 0$, $S = 0$, and $T = 1$. The neutral eta meson η^0 forms an *isospin singlet* with $Y = 0$, $S = 0$, and $T = 0$. The two remaining kaons K^- and $\bar{K}^0$, at last, are seen to form another *isospin doublet* with $Y = -1$, $S = -1$, and $T = 1/2$.

The octet of pseudoscalar mesons is thus composed of one isospin triplet, two isospin doublets, and one isospin singlet. If we denote the (super)multiplets by the general symbol [**i**], where **i** stands for the dimensionality of the representation, then we can write the breaking of the SU(3) symmetry as follows:

$$[\mathbf{8}] \rightarrow [\mathbf{3}] \oplus [\mathbf{2}] \oplus [\mathbf{2}] \oplus [\mathbf{1}]\,. \tag{7.168}$$

7.5.2 Different SU(3) representations

The vector meson octet and baryon octet

When we arrange the vector mesons of Table 7.2 according to their hypercharge Y and isospin component T_3, a similar scheme is obtained as for the pseudoscalar mesons; they form a *vector meson octet* as shown in Figure 7.9.

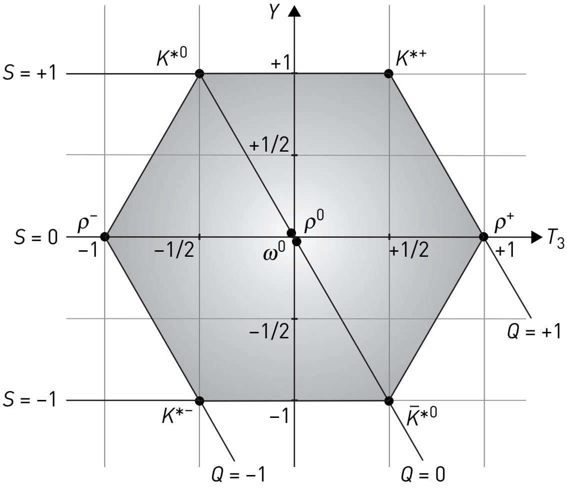


FIGURE 7.9 Weight diagram of the *vector meson octet.*

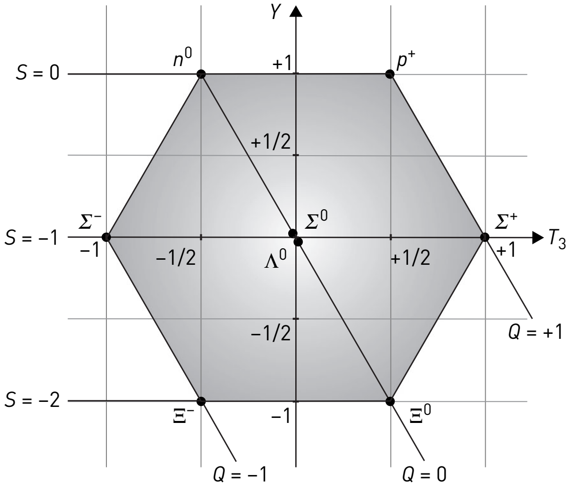


FIGURE 7.10 Weight diagram of the *baryon octet*.

Although the baryons of spin 1/2 (cf. Table 7.3) also fit in the hexagonal pattern, giving rise to a *baryon octet* (Figure 7.10), there are a number of important differences with the two meson octets. First of all, because baryons have baryon number $B = +1$, their strangeness charges are always 1 unit less than the hypercharge, according to Eq. (7.163). For example, the upper line of the baryon octet—with the neutron n^0 and the proton p^+ forming an isospin doublet—has hypercharge $Y = +1$ and strangeness $S = 0$.

Second, the antibaryons are not contained in the baryon octet; they form an octet on their own. This *antibaryon octet* can be obtained from the corresponding baryon octet by placing a bar over each baryon's symbol.

Singlets and decuplets

When classifying the pseudoscalar mesons, Gell-Mann was forced to exclude the eta prime meson η'^0 from inclusion in the octet (cf. Table 7.1). The phi vector meson ϕ^0 likewise was excluded from the vector meson octet (cf. Table 7.2). Instead, both particles were seen to form a multiplet on their own; they gave rise to an SU(3) *singlet*. This is the most trivial SU(3) representation. The baryon resonances of Table 7.4, on the other hand, fit in a triangular scheme forming a *decuplet* (or *decimet*) of particles (Figure 7.11).

Amusingly, the baryon decuplet resembles the *tetractys* (τετρακτύς, or *mystic tetrad*) of the Pythagorean school—a triangular figure with ten points arranged in four rows of 1, 2, 3, and 4 points. This ancient sacred symbol represented the four Platonic elements (water, fire, earth, and air) of which all matter was thought to be composed. We now know that most matter is baryonic, and it is gratifying to see that the baryons fit neatly in the tetractys configuration.

Although hexagons are clearly not the only geometric figures Gell-Mann obtained, not all imaginable shapes are observed. The reason why vector mesons, for instance, form an octet rather than a septet or a nonet of particles can be rationalized only from a deeper understanding of the SU(3) symmetry that underlies their classification. We will study this point in the next section §7.6.

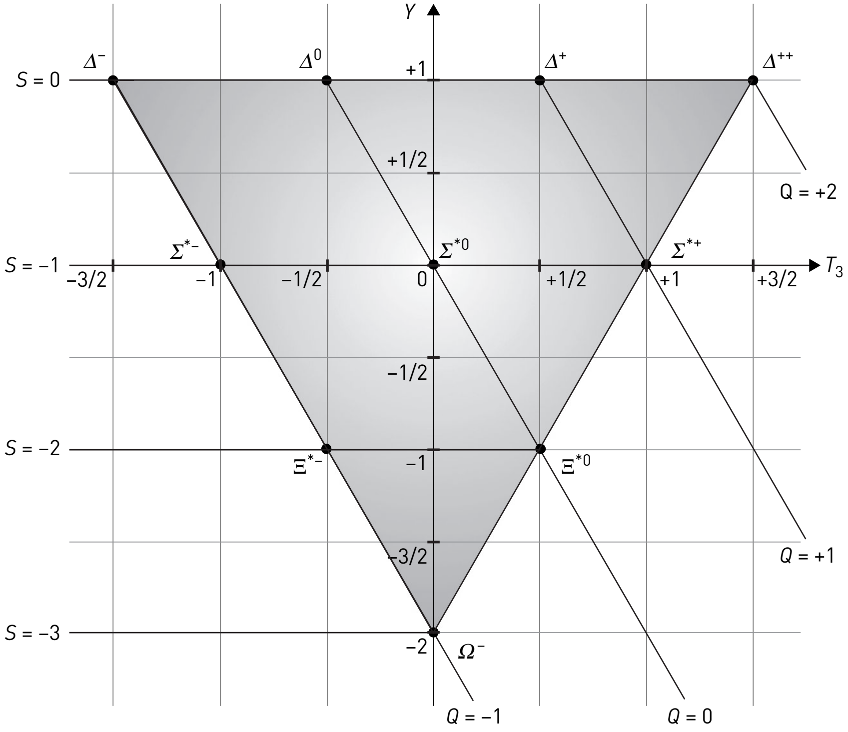


FIGURE 7.11 Weight diagram of the *baryon resonance decuplet.*

7.5.3 Broken symmetries

The Hamiltonian $\hat{\mathscr{H}}_{st}$, which describes the strong nuclear force, is supposed to be *invariant* under the SU(3) group. It therefore commutes with the generators of the $\mathfrak{su}(3)$ Lie algebra:

$$\left[\hat{\mathscr{H}}_{st}, \lambda_i\right] = 0, \qquad \forall i = 1 \rightarrow 8. \tag{7.169}$$

By analogy with the angular momentum algebra (Chapter 5), we can conclude that all the particles of an SU(3) supermultiplet are *degenerate* in their energy. Because energy and mass are related via Einstein's equation $E = mc^2$, we instead speak of a degeneracy in the particle *masses.*

Let us verify to what extent this statement holds by first considering a number of isospin multiplets. The masses of the proton p^+ and the neutron n^0, for example, were given at the beginning of this chapter:

$$\begin{aligned} m_{p^+} &= 938.272046 \text{ MeV}/c^2; \\ m_{n^0} &= 939.565378 \text{ MeV}/c^2. \end{aligned} \tag{7.170}$$

Despite their forming an isospin doublet, a small difference in the mass values is still noticeable, which implies the isospin symmetry is not completely perfect, but slightly *broken.* The same can be said for other isospin multiplets. The three pions, for instance, have the masses as listed in Table 7.1:

$$\begin{aligned} m_{\pi^+} &= 140\ \mathrm{MeV}/c^2; \\ m_{\pi^0} &= 135\ \mathrm{MeV}/c^2; \\ m_{\pi^-} &= 140\ \mathrm{MeV}/c^2. \end{aligned} \tag{7.171}$$

When considering the SU(3) supermultiplets in their totality, even greater differences are observed. The baryon octet shows a mass spread of nearly 400 MeV between its particle masses (cf. Table 7.3); in the pseudoscalar meson octet, the masses of the kaons are nearly three times as high as the pion masses (cf. Table 7.1). Clearly, the SU(3) symmetry is *badly broken*. The reason for this will become clear in §7.6.

The Gell-Mann–Okubo mass formula

Despite these broken symmetries, some deeper logic is obviously at work behind the different particle's masses. The decuplet of baryon resonances, for instance, consists of four isospin multiplets: a Δ quartet $\{\Delta^{++}, \Delta^{+}, \Delta^{0}, \Delta^{-}\}$ with $S = 0$, a Σ^* triplet $\{\Sigma^{*+}, \Sigma^{*0}, \Sigma^{*-}\}$ with $S = -1$, a Ξ^* doublet $\{\Xi^{*0}, \Xi^{*+}\}$ with $S = -2$, and an Ω singlet $\{\Omega^-\}$ with $S = -3$ (Figure 7.11). Their masses are as follows:

$$\begin{aligned} m_{\Delta} &= 1232\ \mathrm{MeV}/c^2; \\ m_{\Sigma^*} &= 1385\ \mathrm{MeV}/c^2; \\ m_{\Xi^*} &= 1530\ \mathrm{MeV}/c^2; \\ m_{\Omega} &= 1672\ \mathrm{MeV}/c^2, \end{aligned} \tag{7.172}$$

with m_{X} standing for the average mass of the X isospin multiplet. It follows from Eq. (7.172) that for every new multiplet (with the strangeness charge decreasing stepwise by 1 unit), the mass raises by roughly 145 MeV/c^2. This led Gell-Mann, in 1961, and Susumu Okubo (b. 1930) independently in 1962, to propose the *Gell-Mann–Okubo mass formula* for the baryon decuplet:

$$m_{\Omega} - m_{\Xi^*} = m_{\Xi^*} - m_{\Sigma^*} = m_{\Sigma^*} - m_{\Delta} \approx 145\ \mathrm{MeV}/c^2. \tag{7.173}$$

Although the Ω^- particle had not yet been observed at the time, Gell-Mann was able to use the previous formula to predict the mass of Ω^-, as recounted in more detail in the next paragraph.

7.5.4 In Mendeleev's footsteps

When Mendeleev first proposed his periodic system in 1869, he consciously left a number of *gaps* in the table—boldly predicting that they would get filled in the future by as-yet-unknown elements. The gaps directly underneath boron, aluminium, and silicon, for instance, hinted at the existence of three new elements that Mendeleev provisionally named *ekaboron* (Eb), *ekaaluminium* (Ea), and *ekasilicon* (Es).[22]

---

[22] Mendeleev adhered to the *Sanskrit numbering system*—using the prefixes *eka-*, *dvi-* and *tri-* for one, two, and three—to denote whether the missing elements were one, two, or three places below a known element in his table. His table was similar to the short form of the periodic table in Figure 13.1A, but with an indent of the periods 1a, 3a, 4b, 5b, 6b. In this way scandium is directly below boron, gallium below aluminium, and germanium below silicon.

News was spreading fast about Mendeleev's daring prophecies, but most chemists remained cautious, if not skeptical; it just seemed a remarkably improbable feat to foresee the existence of unknown elements on the basis of a mere theoretical construct. And as if that wasn't bold enough, Mendeleev went on predicting the chemical and physical properties of the eka elements from their relative position in the periodic system. The atomic weight of ekaboron, for example, could be guessed at by calculating the mean atomic weight of the elements above, below, and to the left and right of Eb.

But, Mendeleev's genius soon became apparent; one by one, and in the most minute details, Mendeleev's prophecies came true! In 1875, French chemist Lecoq de Boisbaudran (1838–1912) discovered *gallium* (Ga) and immediately noted the similarities between the properties of gallium and ekaaluminium.[23] Four years later, Lars Fredrick Nilson (1840–1890) discovered Mendeleev's ekaboron and named it *scandium* (Sc) to honor Scandinavia. Ekasilicon, at last, was discovered in 1882 by German chemist Clemens Winkler (1838–1904), who named it *germanium* (Ge) to honor Germany.

The quest for omega minus

Approximately one hundred years later, Gell-Mann bravely followed in Mendeleev's footsteps, predicting the existence of new particles (along with their properties such as mass, isospin, and strangeness) on the basis of the holes in his schemes. When, at the beginning of 1961, Gell-Mann took on the challenge of organizing the pseudoscalar mesons, only three pions and four kaons were known. With a totality of seven particles, one more particle was clearly missing to complete the octet. Gell-Mann thus predicted the existence of a new meson and dubbed it the *eta particle*, η^0; it was discovered at the Bevatron before the end of 1961.

A similar situation occurred in 1962, when Gell-Mann set out to order the extremely short-lived baryon resonances. The isospin quartet of Δ resonances, triplet of Σ^* resonances, and doublet of Ξ^* resonances, were already known to particle physicists and fitted nicely into a triangular scheme (Figure 7.11). One member was still missing, however, to complete the decuplet of particles. Once again, Gell-Mann proposed the existence of a new particle and called it the *omega minus* baryon Ω^-. From its position in his triangular scheme, Gell-Mann deduced the properties of the Ω^- particle; it would have an isospin charge $T_3 = 0$, hypercharge $Y = -2$, strangeness $S = -3$, and electric charge $Q = -1$. With the help of the Gell-Mann–Okubo mass formula (7.172), the mass of Ω^- was determined to be about 1675 MeV/c^2. Furthermore, Gell-Mann believed the Ω^- particle would be longer lived than the other resonances of the decimet, inasmuch as there were no other (lighter) particles with strangeness $S = -3$ into which the Ω^- particle might decay. As a consequence, strangeness conservation would be violated during the Ω^- decay, and such processes

---

[23] When the news about de Boisbaudran's discovery reached Mendeleev via the French *Comptes Rendus*, Mendeleev immediately noted a difference in his predicted value of the density of ekaaluminium and Lecoq's experimentally determined value. Without hesitation, Mendeleev scribbled a note to de Boisbaudran, urging him to repeat the experiment and to verify the density of gallium. Amused by Mendeleev's faith in his own predictions, de Boisbaurdan followed his advice and proved Mendeleev entirely correct.

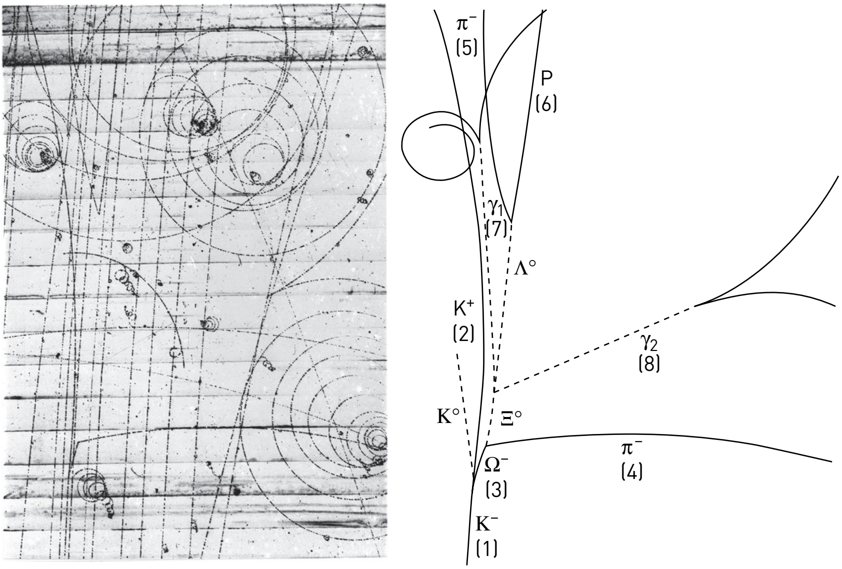


FIGURE 7.12 The discovery of the Ω^- particle. The photograph from the hydrogen bubble chamber is shown on the left; the most relevant tracks are shown in a line diagram on the right. You can see the Ω^- particle decaying into $\Xi^0 + \pi^-$, with subsequent decay of Ξ^0 into $\Lambda^0 + 2\gamma$. (Photo courtesy Brookhaven National Laboratory.)

could happen only via the weak force. This yielded longer lifetimes on the order of 10^{-10} seconds (see §7.1.5).

Well aware of the importance of Gell-Mann's prediction as a crucial test for the eightfold way, particle physicists from every corner of the world embarked on the quest for the tantalizing omega minus. Two years later, in 1964, a particle accelerator group at Brookhaven detected the Ω^- particle in a hydrogen bubble chamber, as illustrated in the photograph in Figure 7.12. With this triumphant discovery, the subatomic puzzle was finally completed and Gell-Mann's oddly shaped schemes were established for good.

7.6 THREE QUARKS FOR MUSTER MARK (1964)

Although Mendeleev had found a way of classifying the elements in his periodic system, he wisely forbore any attempt at explaining the *origin* of the periodic law. The first definite answers to this thorny question would have to await the discovery of the composite structure of the atom (in terms of protons, electrons, and neutrons), and the advent of quantum mechanics and the Pauli exclusion principle during the early 1920s (see Chapter 12).

Soon after the discovery of Ω^-, Gell-Mann similarly started wondering about the possible *origins* of the SU(3) symmetry. Just as the atomic *substructure* had provided an explanation for the periodic system, Gell-Mann believed a true understanding of the eightfold way might be arrived at by postulating that baryons and mesons were *not* elementary particles after all, but had, in fact, a *composite structure*; they were *bound states* of *quarks* and *antiquarks*.

7.6.1 Aces and Quarks

George Zweig (b. 1937) independently the same ideas about the composite structure of the strongly interacting hadrons in 1964, but he called the components *aces* and *deuces*. Gell-Mann, on the other hand, preferred the name *quarks*, which he borrowed from James Joyce's *Finnegans Wake*[24]:

> Three quarks for Muster Mark!
> Sure he has not got much of a bark
> And sure any he has it's all beside the mark.

The number 3 in the previous quotation fitted perfectly because Gell-Mann postulated the existence of three varieties, or *flavors*, of quarks (q): the *up quark* (u), the *down quark* (d), and the *strange quark* (s).[25] For every quark flavor, there is an associated *antiquark* ($\bar{q}$), denoted by a bar over the corresponding quark symbol; we distinguish the *antiup quark* ($\bar{u}$), the *antidown quark* ($\bar{d}$), and the *antistrange quark* ($\bar{s}$). Antiquarks have the same mass as their respective quarks, but the signs of all quantum numbers are reversed (Table 7.10).

Composite particles of quarks and antiquarks are known as *hadrons*. Within a hadron, the quarks and antiquarks are held together via *gluons*, which act like some sort of glue between the quarks.[26] The branch of physics dealing with the quark–quark, quark–gluon, and gluon–gluon interactions is called *quantum chromodynamics* (or QCD for short), for reasons that will become apparent in §7.6.4.

Let us determine some of the intrinsic properties of quarks. To this end, we need to return briefly to our examination of the isotropic harmonic oscillator (§7.3), which also possessed SU(3) symmetry. For each Cartesian coordinate, we introduced a creation operator $a_i^\dagger$ and an annihilation operator a_i, the combinations of which spanned an $\mathfrak{su}(3)$ algebra. In the current context, it is tempting to interpret the $a_i^\dagger$ operators as a set of three independent *quark creation* operators, and the a_i operators as three *quark annihilation* operators (with $i = u, d, s$),[27] which satisfy the commutation relations

$$\left[a_i, a_j^\dagger\right] = I\delta_{ij}. \tag{7.174}$$

The remaining commutators are all equal to zero:

$$\left[a_i, a_j\right] = \left[a_i^\dagger, a_j^\dagger\right] = \left[a_i, I\right] = \left[a_i^\dagger, I\right] = 0. \tag{7.175}$$

---

[24] M. Gell-Mann. *The Quark and the Jaguar*. 1994, p. 180.

[25] In reality, three more flavors of heavy quarks are known to exist: the *charm quark* (c), the *beauty* or *bottom quark* (b), and the *truth* or *top quark* (t) (see §7.6.4).

[26] The hadron constituents (i.e., quarks, antiquarks, and gluons) are known collectively as *partons*, a term first coined by theoretical physicist Richard Feynman (1918–1988).

[27] Fuchs and Schweigert, "The Lie Algebras SU(3) and Hadron Symmetries," p. 40. See also H. Georgi. "Flavor SU(3) Symmetries in Particle Physics." *Physics Today* 41.4 (1988), pp. 29–37. The three-dimensional harmonic oscillator can also be seen as a system of three uncoupled harmonic oscillators in one dimension of the same frequency. As Fuchs and Schweigert observed, the fact "that the oscillators have the same frequency means that the three particles should have equal masses" (p. 40).

Table 7.10 Table of the properties of the low-mass *quark flavors* (up u, down d, and strange s) and their corresponding *antiquarks* (antiup $\bar{u}$, antidown $\bar{d}$, and antistrange $\bar{s}$). All quarks and antiquarks have spin $J = 1/2$; as a result of their half-integer spin, quarks are classified as *fermions*. Antiquarks have the same mass and total isospin T as their respective quarks, but the signs of all flavor quantum numbers T_3, Y, B, S, and Q have been reversed. The baryon number $B = +1/3$ for all quarks, whereas $B = -1/3$ for the antiquarks. As a general rule, the lower the mass of the quark, the more stable it will be. The s quark, therefore, typically decays into the lighter u and d quarks. For this reason, the u and d quarks are the most abundant in the universe, whereas the s quark is only produced in the interior of particle accelerators.

Symbol	*Isospin (T)*	*Isospin Component (T_3)*	*Hyper-charge (Y)*	*Baryon Number (B)*	*Strangeness (S)*	*Charge (Q)*	*Rest Mass (MeV/c^2)*
u	1/2	+1/2	+1/3	+1/3	0	+2/3	1.7–3.3
d	1/2	−1/2	+1/3	+1/3	0	−1/3	4.1–5.8
s	0	0	−2/3	+1/3	−1	−1/3	101
$\bar{u}$	1/2	−1/2	−1/3	−1/3	0	−2/3	1.7–3.3
$\bar{d}$	1/2	+1/2	−1/3	−1/3	0	+1/3	4.1–5.8
$\bar{s}$	0	0	+2/3	−1/3	+1	+1/3	101

We consider the three quark flavors as resulting from the action of the creation operators $a_u^\dagger$, $a_d^\dagger$, and $a_s^\dagger$ on the *vacuum state* $|0\rangle$:

$$|u\rangle = a_u^\dagger\,|0\rangle\,; \quad |d\rangle = a_d^\dagger\,|0\rangle\,; \quad |s\rangle = a_s^\dagger\,|0\rangle\,. \tag{7.176}$$

Similarly, acting on any of these quark states with the annihilation operators a_u, a_d, and a_s destroys the quark state and reduces it back to the vacuum state:

$$|0\rangle = a_u\,|u\rangle\,; \quad |0\rangle = a_d\,|d\rangle\,; \quad |0\rangle = a_s\,|s\rangle. \tag{7.177}$$

To deduce the hypercharge Y and the isospin component T_3 of these quarks, we have to act on them with the corresponding hypercharge operator $\hat{Y}$ and isospin component operator $\hat{T}_3$, defined in Eq. (7.138), §7.4.6:

$$\hat{Y} \equiv \frac{1}{\sqrt{3}}\lambda_8; \quad \hat{T}_3 \equiv \frac{1}{2}\lambda_3. \tag{7.178}$$

From the matrix form of the Gell-Mann matrices (Table 7.5), we obtain

$$\begin{aligned}\hat{Y} &= \frac{1}{\sqrt{3}}\begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & -\frac{2}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{bmatrix}; \\ \hat{T}_3 &= \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned} \tag{7.179}$$

Having found the matrix representation of $\hat{Y}$ and $\hat{T}_3$, we now have to translate this in operator form. It is possible to construct an operator Lie algebra isomorphic to a matrix Lie algebra by associating each $n \times n$ matrix $\mathbb{A}$ of the matrix algebra to an operator $\hat{A}$ of the operator algebra that is a linear combination of the n creation and n annihilation operators:

$$\mathbb{A} \to \hat{A} = \mathbf{a}^\dagger \mathbb{A}\mathbf{a} = \sum_i \sum_j a_i^\dagger \mathbb{A}_{ij} a_j. \tag{7.180}$$

We thus obtain for the $\hat{Y}$ operator

$$\begin{aligned}\hat{Y} &= \begin{bmatrix} a_u^\dagger & a_d^\dagger & a_s^\dagger \end{bmatrix}\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{bmatrix}\begin{bmatrix} a_u \\ a_d \\ a_s \end{bmatrix} \\ &= \frac{1}{3}a_u^\dagger a_u + \frac{1}{3}a_d^\dagger a_d - \frac{2}{3}a_s^\dagger a_s.\end{aligned} \tag{7.181}$$

Similarly for $\hat{T}_3$, we get

$$\begin{aligned}\hat{Y} &= \begin{bmatrix} a_u^\dagger & a_d^\dagger & a_s^\dagger \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_u \\ a_d \\ a_s \end{bmatrix} \\ &= \frac{1}{2}a_u^\dagger a_u - \frac{1}{2}a_d^\dagger a_d.\end{aligned} \tag{7.182}$$

Acting with $\hat{Y}$ on $|u\rangle$ then yields

$$\hat{Y}\,|u\rangle = \left(\frac{1}{3}a_u^\dagger a_u + \frac{1}{3}a_d^\dagger a_d - \frac{2}{3}a_s^\dagger a_s\right)|u\rangle = \frac{1}{3}\,|u\rangle. \tag{7.183}$$

It follows that the hypercharge $Y = 1/3$. Similarly, acting with $\hat{T}_3$ on $|u\rangle$ yields

$$\hat{T}_3\,|u\rangle = \left(\frac{1}{2}a_u^\dagger a_u - \frac{1}{2}a_d^\dagger a_d\right)|u\rangle = \frac{1}{2}\,|u\rangle, \tag{7.184}$$

with $T_3 = +1/2$ for the value of the isospin component.

We can proceed in the same manner for the down quark:

$$\begin{aligned}\hat{Y}\,|d\rangle &= +\frac{1}{3}\,|d\rangle;\\ \hat{T}_3\,|d\rangle &= -\frac{1}{2}\,|d\rangle,\end{aligned} \tag{7.185}$$

yielding $Y = +1/3$ and $T_3 = -1/2$. Finally, for the strange quark

$$\begin{aligned}\hat{Y}\,|s\rangle &= -\frac{2}{3}\,|s\rangle;\\ \hat{T}_3\,|s\rangle &= \;\;0\,|s\rangle,\end{aligned} \tag{7.186}$$

with $Y = -2/3$ and $T_3 = 0$. The u, d, and s quark can now be written as

$$|u\rangle = |T_3, Y\rangle = \left|+\frac{1}{2}, \frac{1}{3}\right\rangle; \tag{7.187}$$

$$|d\rangle = |T_3, Y\rangle = \left|-\frac{1}{2}, \frac{1}{3}\right\rangle; \tag{7.188}$$

$$|s\rangle = |T_3, Y\rangle = \left|0, -\frac{2}{3}\right\rangle. \tag{7.189}$$

The three quark flavors are assigned a fractional baryon number $B = 1/3$. Following Eq. (7.163), this results in a strangeness charge of $S = 0$ for the u and d quark, and $S = -1$ for the s quark. Clearly, the strangeness quantum number is just a measure for the amount of strange quarks in any quantum system. The electric charges of the three quarks can be deduced via the Gell-Mann–Nishijima charge formula (7.164):

$$Q_u = T_3 + \frac{1}{2}Y = +\frac{2}{3}; \tag{7.190}$$

$$Q_d = T_3 + \frac{1}{2}Y = -\frac{1}{3}; \tag{7.191}$$

$$Q_s = T_3 + \frac{1}{2}Y = -\frac{1}{3}. \tag{7.192}$$

The up quark has $Q = +2/3$, whereas the down and strange quarks share the same charge number $Q = -1/3$. All these values are listed in Table 7.10.

When the three quark flavors are plotted in terms of their Y and T_3 eigenvalues in the Y—T_3-plane, a triangular eightfold way pattern is obtained (Figure 7.13). This triplet of quarks forms the *fundamental representation* of the SU(3) group (i.e., the smallest nontrivial SU(3) representation). In the respective antiquark multiplet, the signs of the additive quantum numbers T_3, Y, S, and Q are reversed, as seen on the right-hand side of Figure 7.13 (see also Table 7.10). Both the *quark* and the *antiquark multiplets* contain an isospin doublet with $T = 1/2$ and an isospin singlet with $T = 0$. The isodoublet of the quark triplet, for instance, consists of the u and d quark, whereas the s quark falls into an isosinglet. The same applies to the

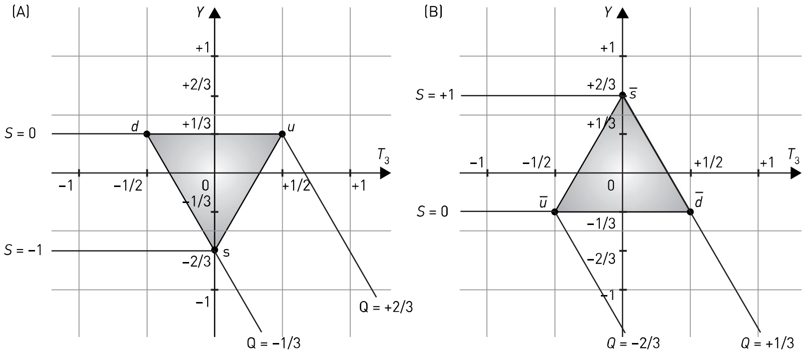


FIGURE 7.13 Weight diagram of the *quark triplet* (A) and the *antiquark triplet* (B).

three antiquarks: $\bar{u}$ and $\bar{d}$ forming an isodoublet and $\bar{s}$, an isosinglet. As we will see later, all higher dimensional meson and baryon multiplets can be built from these two fundamental triplets (see §7.6.2 and §7.6.3).

We conclude this section with a brief consideration of the $\mathfrak{su}(3)$ Weyl generators which act as shift operators on the u, d, and s quarks, transforming them into one another. We proceed in analogy with the $\hat{Y}$ and $\hat{T}_3$ operators to deduce the matrix form of the Weyl operators defined in Eq. (7.142). The $\hat{V}_+$ operator, for instance, has the following form:

$$\begin{aligned}\hat{V}_+ = \frac{1}{2}(\lambda_4 + i\lambda_5) &= \frac{1}{2}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + i\frac{1}{2}\begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned} \tag{7.193}$$

In operator form, this becomes

$$\hat{V}_+ = \begin{bmatrix} a_u^\dagger & a_d^\dagger & a_s^\dagger \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_u \\ a_d \\ a_s \end{bmatrix} = a_u^\dagger a_s. \tag{7.194}$$

Acting with $\hat{V}_+$ on the strange quark $|s\rangle$ yields

$$\hat{V}_+ |s\rangle = a_u^\dagger a_s |s\rangle = a_u^\dagger |0\rangle = |u\rangle. \tag{7.195}$$

The strange quark has been transformed in an up quark under the influence of $\hat{V}_+$. Figure 7.14 shows the action of the six Weyl operators on the fundamental representation of the SU(3) group.

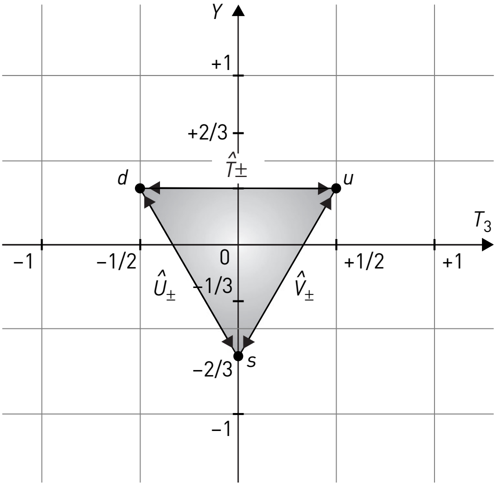


FIGURE 7.14 Action of the six Weyl operators on the fundamental representation of the SU(3) group.

Table 7.11 Table of the $q\bar{q}$ combinations, along with their total charge Q, isospin component T_3, strangeness S, and hypercharge Y. The resulting mesons are identified in the last two columns. *Pseudoscalar mesons* are formed when the spins of the quark and antiquark are pointing in opposite directions; *vector mesons* are obtained when the q and $\bar{q}$ spins are aligned in the same direction.

$q\bar{q}$	Q	T_3	S	Y	*Pseudoscalar Meson*	*Vector Meson*
$u\bar{u}$	0	0	0	0	π^0	ρ^0
$u\bar{d}$	+1	+1	0	0	π^+	ρ^+
$u\bar{s}$	+1	+1/2	+1	+1	K^+	K^{*+}
$d\bar{u}$	−1	−1	0	0	π^-	ρ^-
$d\bar{d}$	0	0	0	0	η^0	ω^0
$d\bar{s}$	0	−1/2	+1	+1	K^0	K^{*0}
$s\bar{u}$	−1	−1/2	−1	−1	K^-	K^{*-}
$s\bar{d}$	0	+1/2	−1	−1	$\bar{K}^0$	$\bar{K}^{*0}$
$s\bar{s}$	0	0	0	0	η'^0	ϕ^0

7.6.2 Mesons

Mesons are defined as *bound states* of a quark and an antiquark; they form $q\bar{q}$ pairs.[28] The nine possible $q\bar{q}$ combinations are enumerated in Table 7.11, along with their total charge Q, strangeness S, hypercharge Y, and isospin component T_3. Note that because Q, S, Y, and T_3 are *additive* quantum numbers, the values listed in Table 7.11 are simply the *sum* of the q and $\bar{q}$ quantum numbers. For example, the combination of an u quark with $Q = +2/3$ and a $\bar{d}$ antiquark with $Q = +1/3$ yields a $u\bar{d}$ meson with a total charge of $Q = +1$.

---

[28] At the beginning of this chapter (§7.1.2), mesons were defined as *middleweight* particles with a mass between those of baryons and leptons. Considering that some mesons were discovered later that are *heavier* than baryons (e.g., $m_{\eta'^0} > m_{p^+}$ and $m_{\phi^0} > m_{n^0}$), the definition of mesons as $q\bar{q}$ pairs turns out to be more correct.

In the same vein, we find that the baryon number B of mesons always equals zero, because quarks have baryon number $B = +1/3$ and antiquarks have $B = -1/3$. Quarks and antiquarks have spin $J = 1/2$; when their spins are pointing in *opposite* directions, the total spin is zero and a *pseudoscalar meson* is formed. When both spins point in the *same* direction, the total spin $J = 1$ and a *vector meson* is obtained.

Comparing the values for Q, S, Y, and T_3 in Table 7.11 with those of the pseudoscalar mesons and vector mesons in Tables 7.1 and 7.2, respectively, helps to identify the $q\bar{q}$ combinations, as shown in Table 7.11.

The process of meson formation out of quarks and antiquarks can also be represented in a diagrammatic way. To this aim, we can form the *direct product* (or *tensor product*) of a quark triplet with an antiquark triplet, and expect a hadron multiplet to result. We use the same graphical method for forming product representations as outlined in Chapter 6 (§6.4), and superimpose a triplet of antiquarks on top of each member state of the quark triplet, taking care that the center of gravity of the antiquark triplet coincides with every site of the quark triplet. This is illustrated in Figure 7.15.

In the end, nine meson states are formed; six states lie on the corners of a regular hexagon and the three remaining states are located in the center of the graph. This composite system can be decomposed into a *direct sum* of a meson octet and a meson singlet. That is, although the eight states of the meson octet are seen to transform among themselves under the operations of the SU(3) group, they do *not* mix with the singlet state (which we identity as the η'^0 pseudoscalar meson or the ϕ^0 vector meson). We can write this process symbolically as

$$[\mathbf{3}] \otimes \left[\bar{\mathbf{3}}\right] = [\mathbf{8}] \oplus [\mathbf{1}], \tag{7.196}$$

where $[\mathbf{3}]$ represents the quark triplet and $\left[\bar{\mathbf{3}}\right]$ stands for the antiquark triplet. Note that we use the dimensions of the multiplets to label the irreducible representations of the SU(3) group.

The three $q\bar{q}$ states in the center of the graph ($u\bar{u}$, $d\bar{d}$, and $s\bar{s}$) all have $T_3 = 0$ and $Y = 0$. It is therefore impossible to say which of these corresponds to the π^0 meson, which is the η^0 meson, and which the η'^0 meson. It turns out that the π^0, η^0, and η'^0 mesons are *mixtures* of the $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ states. However interesting a phenomenon, we will not pursue this line of reasoning any further.

We conclude this section by noting that the *antimesons* are readily obtained from their corresponding *mesons* by turning every quark into an antiquark and vice versa. The antiparticle of π^- (with a quark content $d\bar{u}$), for instance, is π^+ (with the quark content $u\bar{d}$).

7.6.3 Baryons

Baryons are defined as *bound states* of three quarks, denoted abstractly as qqq. A proton, for example, consists of two u quarks and one d quark embedded in a glob of gluons that holds them together. The ten possible qqq combinations are enumerated in Table 7.12, along with their total charge Q, strangeness S, hypercharge Y, and isospin component T_3. *Antibaryons* are similarly defined as bound states of three antiquarks, denoted $\bar{q}\bar{q}\bar{q}$. Because quarks are assigned a baryon number of $B = +1/3$, all baryons have $B = +1$. Antibaryons, by analogy, have $B = -1$ because antiquarks always have $B = -1/3$.

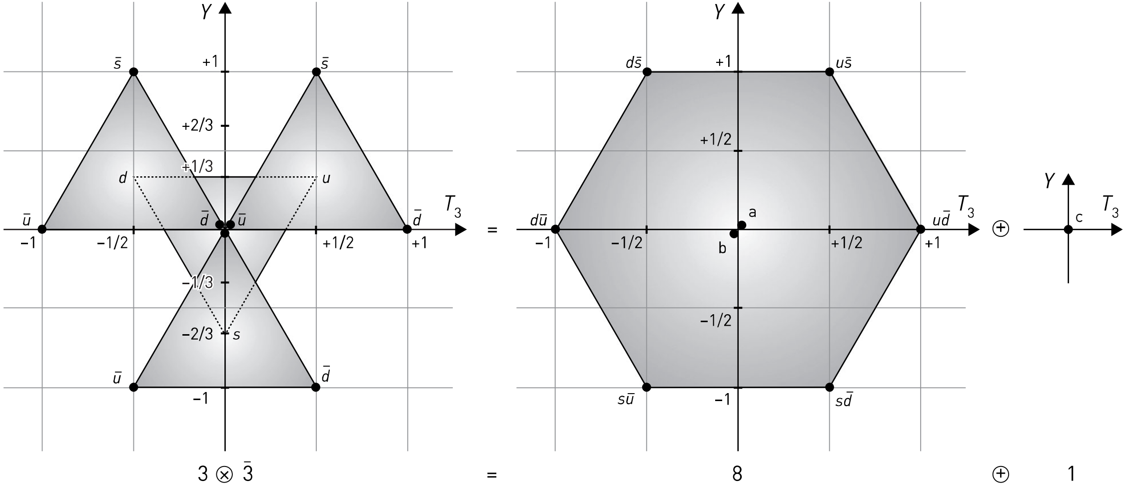


FIGURE 7.15 Weight diagrams illustrating the reduction of the tensor product $[\mathbf{3}] \otimes [\bar{\mathbf{3}}] = [\mathbf{8}] \oplus [\mathbf{1}]$, yielding the $q\bar{q}$ SU(3) multiplets. As a result of quantum mechanical mixing, the quark content of the points labeled *a*, *b*, and *c* has not been assigned.

Table 7.12 Table of the qqq combinations, along with their total charge Q, isospin component T_3, strangeness S, and hypercharge Y. The resulting *baryon resonances* with $B=+1$ and $J=+3/2$ are identified in the last column. The corresponding table of $\bar{q}\bar{q}\bar{q}$ combinations can be readily obtained by putting a bar over each q symbol, and by inverting the signs of every Q, T_3, S, and Y quantum number.

qqq	Q	T_3	S	Y	*Baryon Resonance*
uuu	$+2$	$+3/2$	0	$+1$	Δ^{++}
uud	$+1$	$+1/2$	0	$+1$	Δ^{+}
udd	0	$-1/2$	0	$+1$	Δ^{0}
ddd	-1	$-3/2$	0	$+1$	Δ^{-}
uus	$+1$	$+1$	-1	0	Σ^{*+}
uds	0	0	-1	0	Σ^{*0}
dds	-1	-1	-1	0	Σ^{*-}
uss	0	$+1/2$	-2	-1	Ξ^{*0}
dss	-1	$-1/2$	-2	-1	Ξ^{*-}
sss	-1	0	-3	-2	Ω^{-}

When the spins of the three quarks are aligned, the total spin equals $J=+3/2$, which corresponds to the spin of the *baryon resonances* (cf. Table 7.4). By comparing the values for Q, S, Y, and T_3 in the Tables 7.12 and 7.4, we can bring the various qqq combinations into correspondence with the ten baryon resonances, as summarized in Table 7.12. (The corresponding antiparticles are once again obtained by turning every quark into an antiquark.)

Let us end this section by determining the process of baryon formation out of three quarks. To do this, we have to form the *direct product* of three quark triplets, and see which hadron multiplets result from the process. Because the reduction of the resulting product representation is more involved than that for mesons, let us simplify the process by first combining two quarks. After superimposing a quark triplet on top of each state of another quark triplet, a total of nine qq combinations are obtained that fall into a triangular array, as illustrated in Figure 7.16.

Notice that the sites on the corners of the triangle are occupied *singly* whereas the others are occupied *doubly*. We therefore reduce the product representation to a direct sum of a *sextet* and an *antitriplet*, which we recognize as the fundamental antiquark triplet. Symbolically, this is written as

$$[\mathbf{3}] \otimes [\mathbf{3}] = [\mathbf{6}] \oplus \left[\bar{\mathbf{3}}\right]. \tag{7.197}$$

We are now in a position to add the third and final quark to the qq combinations. Using Eq. (7.197), this yields

$$\begin{aligned} [\mathbf{3}] \otimes [\mathbf{3}] \otimes [\mathbf{3}] &= [\mathbf{3}] \otimes \left([\mathbf{6}] \oplus \left[\bar{\mathbf{3}}\right]\right) \\ &= \left([\mathbf{3}] \otimes [\mathbf{6}]\right) \oplus \left([\mathbf{3}] \otimes \left[\bar{\mathbf{3}}\right]\right). \end{aligned} \tag{7.198}$$

We already know from Eq. (7.196) that the direct product of a quark triplet and an antiquark triplet results in an octet and a singlet. It thus remains to be investigated

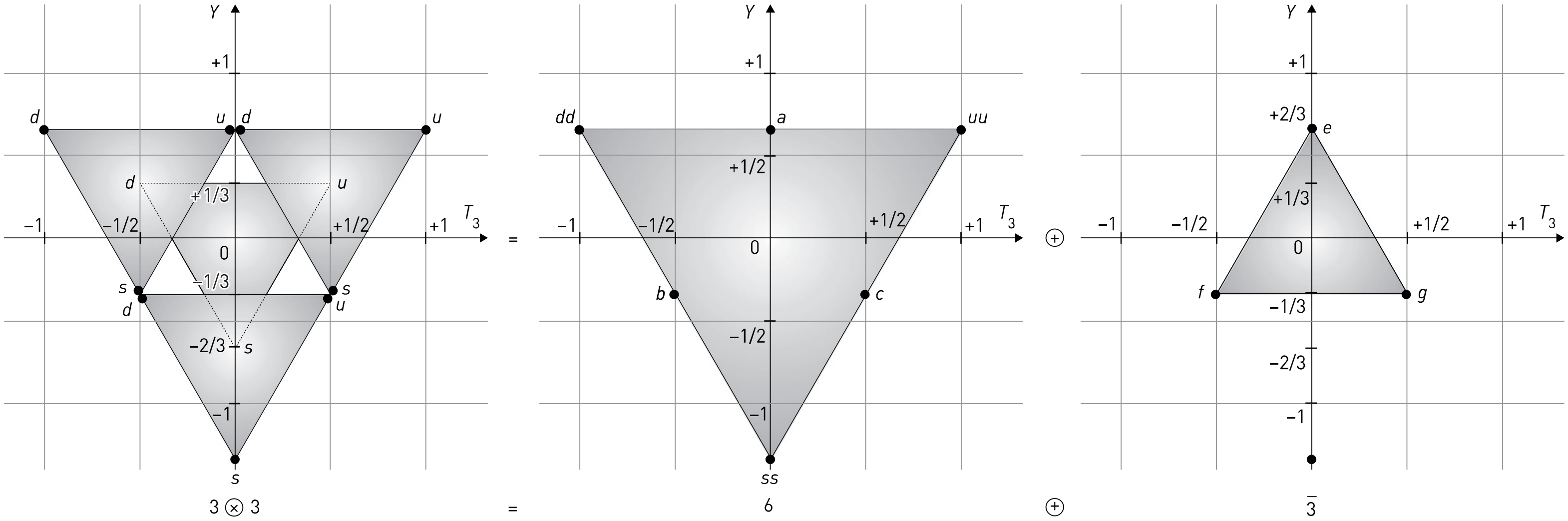


FIGURE 7.16 Weight diagrams illustrating the reduction of the tensor product $[\mathbf{3}] \otimes [\mathbf{3}] = [\mathbf{6}] \oplus [\bar{\mathbf{3}}]$, yielding *qq* SU(3) multiplets. As a result of quantum mechanical mixing, the quark content of the points labeled *a*, *b*, *c*, *e*, *f*, and *g* has not been assigned.

which SU(3) multiplets are formed during the process of adding a third quark to the sextet, which is illustrated in Figure 7.17. The scheme of points thus obtained can be decomposed into a decuplet and an octet:

$$[\mathbf{3}] \otimes [\mathbf{6}] = [\mathbf{10}] \oplus [\mathbf{8}]. \qquad (7.199)$$

We then obtain the final equation:

$$[\mathbf{3}] \otimes [\mathbf{3}] \otimes [\mathbf{3}] = [\mathbf{10}] \oplus [\mathbf{8}] \oplus [\mathbf{8}] \oplus [\mathbf{1}]. \qquad (7.200)$$

The assignment of the quark content to every hadron from the baryon octets and decuplet is gravely complicated by the quantum mechanical mixing of quarks and antiquarks, and will not be pursued here.

7.6.4 Symmetry from the quark perspective

It follows from Tables 7.11 and 7.12 that the strangeness charge S is a measure for the number of strange quarks in hadrons. Every strange quark s imparts a strangeness $S = -1$ to the quantum system, whereas every strange antiquark $\bar{s}$ yields $S = +1$.

Another observation is that isospin is the symmetry under the *interchange* of the u and d quarks. Because the u and d quark have nearly the same mass (cf. Table 7.10), isospin symmetry is a good symmetry. This also explains the near equality of the proton and neutron masses.

Similarly, SU(3) turns out to be the symmetry under the *interchange* of the three flavors of quarks u, d, and s; it is therefore termed *flavor* SU(3). If the three quarks had identical masses, flavor SU(3) would be an exact symmetry. However, the s quark is nearly 100 MeV/c^2 heavier than the u and d quarks (cf. Table 7.10). Kaons, which are known to contain an s quark or $\bar{s}$ antiquark (cf. Table 7.11), are, for that reason, nearly four times as heavy as the π-mesons (see Table 7.1). As a consequence, the SU(3) symmetry is badly broken. One says that the symmetry of the eightfold way is broken by the mass differences between the u, d, and s quarks.

Realism versus instrumentalism: To be or not to be?

The power of the *quark hypothesis* has been demonstrated amply in the previous sections. Yet, most physicists were at first reluctant to adopt Gell-Mann's attractive conjecture. During the 1960s and early 1970s, widespread skepticism reigned within the physics community. For one thing, the *fractional* charges of the u, d, and s quarks (Table 7.10) challenged the preconceived views of electric charge as occuring in *integral* units of the elementary charge e.

Another embarrassing fact was the following: despite the efforts of particle physicists, no one had ever succeeded to produce a *free*, *individual* quark in the heart of particle accelerators. However hard a hadron was hit, its constituent quarks did not bob up. This conveyed the impression that the quarks were somehow imprisoned in the baryons and mesons—a notion referred to as *quark confinement*.

The reasoning behind this ran along the following lines: quarks are held together in hadrons by *gluons*. Whenever someone tries to separate a bound state of two quarks—that is, whenever one attempts to rip a meson apart—gluon matter builds up in between the quarks and acts as some sort of glue holding the quarks together even more tightly. The further we try to separate the two quarks, the more energy we

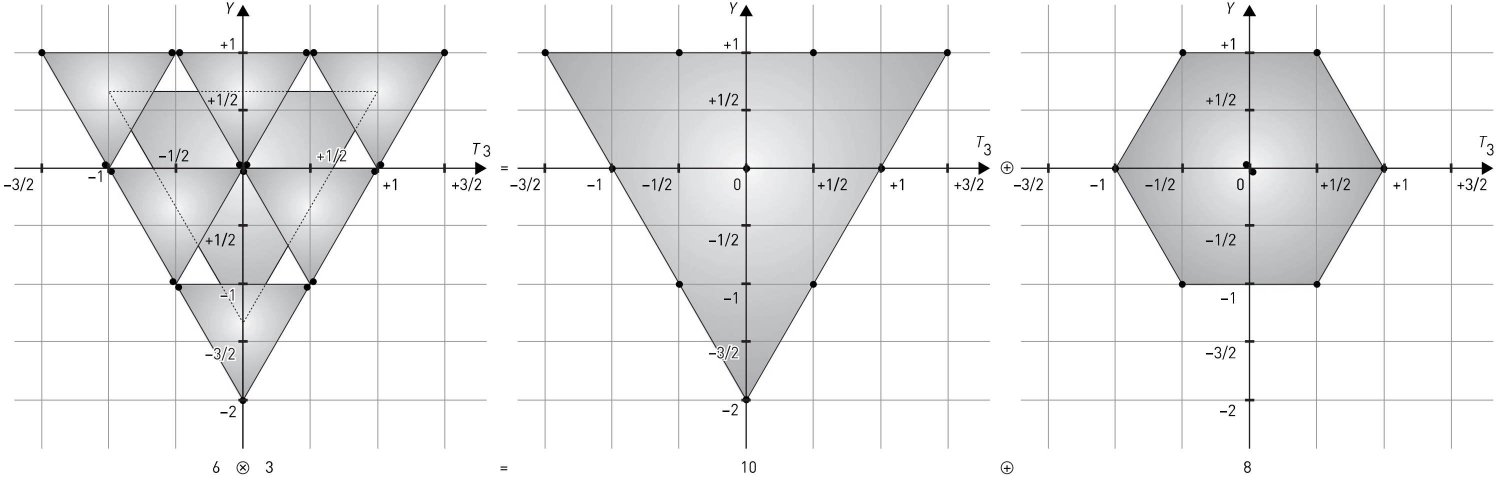


FIGURE 7.17 Weight diagrams illustrating the reduction of the tensor product [**6**] ⊗ [**3**] = [**10**] ⊕ [**8**]. As a result of quantum mechanical mixing, the quark content has not been indicated in this figure.

have to pump into the system. Clearly, the forces acting on the quark level are of a completely different nature than, say, Coulomb's electromagnetic force or Newton's gravitational force, where both are known to *diminish* quadratically with distance, rather than *increase* indefinitely as in the quark case.

The mysteries surrounding the phenomenon of quark confinement raised the question whether quarks were actually *real*. Instead of adopting a *realist* point of view, perhaps quarks should be considered as mere theoretical constructs that could be used only in an *instrumentalist* way to rationalize the symmetries of the eightfold way? However, during the late 1960s, deep inelastic scattering experiments were performed at the Stanford Linear Accelerator Center (SLAC) to probe the interior of the proton. The experiments revealed that the proton contained three smaller pointlike particles, thus adding further support to the quark model.

Flavor SU(3) versus color SU(3)

Given that quarks are fermions (as a result of their half-integer spin of $J = 1/2$), they must be subject to Pauli's *exclusion principle*. That is, no two identical quarks can occupy the same quantum state simultaneously in a hadron. Oddly enough, three s quarks (in identical quantum states) were known to coexist inside the Ω^- particle ($\Omega^- = sss$; cf. Table 7.12).[29] How was this possible without violating Pauli's principle? As a way out of this problem, American physicist Oscar Greenberg (b. 1932) postulated the existence of yet another quantum number in 1964; quarks were said to possess *color*.[30]

Each flavor of quarks (u, d, or s) was believed to occur in three color varieties: *red* (r), *green* (g), or *blue* (b). The corresponding antiquarks ($\bar{u}$, $\bar{d}$, or $\bar{s}$) were colored *antired* ($\bar{r}$), *antigreen* ($\bar{g}$), or *antiblue* ($\bar{b}$). When forming the Ω^- particle (or any other baryon for that matter), one simply had to combine one red, one green, and one blue quark. Then, the three s quarks in the Ω^- particle would all be colored differently and would therefore be distinguishable ($\Omega^- = s_r s_g s_b$); Pauli's principle would no longer be violated.

Quarks are, of course, not *really* colored, but the color terminology offered an interesting analogy to the field of optics. Back in the 1670s, Newton demonstrated how to combine light of three primary colors to obtain a beam of white (i.e., colorless) light. In complete analogy, the r, g, and b quarks were thought to combine to form a "*colorless*" baryon. Mesons, which are $q\bar{q}$ combinations, could be made colorless as well by combining an r (g or b) quark with a $\bar{r}$ ($\bar{g}$ or $\bar{b}$) antiquark. Finally, a combination of three antiquarks, each with a different color charge, would result in a white antibaryon. Greenberg therefore advanced the idea that all naturally occuring hadrons were *color-neutral* particles. Note that this also offered a plausible explanation for the orgin of the quark confinement.

Recall that the three quark flavors (u, d, and s) were seen to form a flavor triplet transforming under the fundamental representation of the flavor SU(3) group (or $\mathrm{SU}(3)_f$ for short); correspondingly, the three color charges (r, g, and b) now formed a *color triplet* transforming under the fundamental representation of the *color* SU(3) group (abbreviated $\mathrm{SU}(3)_c$). However, although both systems are described by the

---

[29] In the same vein, the Δ^{++} consists of three u quarks, and the Δ^- consists of three d quarks.

[30] Hence the name quantum *chromo*dynamics for the theory describing the strong force.

Table 7.13 Table of the properties of the high-mass *quark flavors* (charm c, bottom or beauty b, and top or truth t) and their corresponding *antiquarks* (anticharm $\bar{c}$, antibottom $\bar{b}$, and antitop $\bar{t}$). All quarks and antiquarks have spin $J = 1/2$; because of their half-integer spin, quarks are classified as *fermions*. The total isospin $T = 0$ for c, b, and t. The symbols C, B', and T' denote the quantum numbers *charm*, *bottomness*, and *topness*. Antiquarks have the same mass as their respective quarks, but the signs of all the flavor quantum numbers B, C, B', T', and Q are reversed. The baryon number $B = +1/3$ for all quarks whereas $B = -1/3$ for the antiquarks. The indicated rest masses are only approximate. Note that the top mass is nearly as great as a gold atom; the high-mass quarks are therefore highly unstable and can only be produced in particle accelerators.

Symbol	Baryon Number (B)	*Charm (C)*	*Bottomness (B')*	*Topness (T')*	*Charge (Q)*	*Rest Mass (MeV/c^2)*
c	$+1/3$	$+1$	0	0	$+2/3$	± 1270
b	$+1/3$	0	-1	0	$-1/3$	± 4190
t	$+1/3$	0	0	$+1$	$+2/3$	$\pm 172{,}000$
$\bar{c}$	$-1/3$	-1	0	0	$-2/3$	± 1270
$\bar{b}$	$-1/3$	0	$+1$	0	$+1/3$	± 4190
$\bar{t}$	$-1/3$	0	0	-1	$-2/3$	$\pm 172{,}000$

same SU(3) symmetry, flavor SU(3) is an approximate symmetry (*vide supra*) whereas color SU(3) is believed to be an exact symmetry.

SU(4) and beyond

With the introduction of the other quark flavors—*charm* (c) by Sheldon Lee Glashow, John Iliopoulos, and Luciano Maiani in 1970, and *beauty* (b) and *truth* (t) by Makoto Kobayashi and Toshihide Maskawa in 1973—even higher symmetries were searched for to accommodate these quarks in a unifying scheme (see Table 7.13). Models based on an SU(4), SU(5), or SU(6) symmetry were proposed in the years following Gell-Mann's pioneering work. However, for $n > 3$ (with n being the number of quark flavors), the flavor SU(n) symmetry is increasingly broken. The reason for this can, once again, be traced to the masses of the quarks: the c, b, and t quarks are significantly heavier than the u, d, and s quarks, with the top quark being nearly as heavy as a gold atom (cf. Table 7.13).

8 SU(2) and electron spin

Thus, we arrive at the conclusion, that the Dirac particle and the quantum mechanical rotor are identical dynamic systems. In other words: a Dirac particle is neither more nor less than a particle, for which it is possible to talk about an orientation in space.
–Jens Peder Dahl (1977)[1]

8.1 FROM SU(3) TO SU(2)

In the previous chapter, we saw that SO(3) is a subgroup of SU(3). The subduction was based on the angular momentum operators, which are common to both groups. There is, however, another more subtle way in which rotational symmetry can be retrieved in SU(3): a path via the subgroup SU(2), which takes us to the quantum mechanics of the spinning electron. The path starts by reducing SU(3) to SU(2). This simply consists of taking out one of the degrees of freedom. By removing the a_z and $a_z^\dagger$ operators, only three of the eight $\hat{O}(\lambda)$ operators survive:

$$\begin{aligned}
\hat{O}(\lambda_1) &= a_x^\dagger a_y + a_y^\dagger a_x; \\
\hat{O}(\lambda_2) &= \frac{1}{i}\left(a_x^\dagger a_y - a_y^\dagger a_x\right); \\
\hat{O}(\lambda_3) &= a_x^\dagger a_x - a_y^\dagger a_y.
\end{aligned} \tag{8.1}$$

---

[1] From J. P. Dahl. “The spinning electron.” *Det Kongelige Danske Videnskabernes Selskab Mathematisk-fysiske Meddelelser*, 39.12 (1977), pp. 1–33. See: http://gymarkiv.sdu.dk/MFM/kdvs/mfm%2030-39/mfm-39-12.pdf.

We can think of these operators as acting in an abstract two-dimensional space that allows complex transformations. The $\hat{O}(\lambda)$ operators are closed under commutation:

$$\begin{aligned} \left[\hat{O}(\lambda_1), \hat{O}(\lambda_2)\right] &= 2i\hat{O}(\lambda_3), \\ \left[\hat{O}(\lambda_2), \hat{O}(\lambda_3)\right] &= 2i\hat{O}(\lambda_1), \\ \left[\hat{O}(\lambda_3), \hat{O}(\lambda_1)\right] &= 2i\hat{O}(\lambda_2), \end{aligned} \tag{8.2}$$

which may be rewritten in a concise form by using the Levi-Civita symbol

$$\left[\hat{O}_k, \hat{O}_l\right] = 2i\epsilon_{klm}\hat{O}_m. \tag{8.3}$$

Here k, l, and m are permutations of the values 1, 2, and 3. The rule is entirely similar to the commutation rule for angular momenta in units of $\hbar$:

$$\left[\hat{L}_k, \hat{L}_l\right] = i\epsilon_{klm}\hat{L}_m, \tag{8.4}$$

where k, l, and m stand for permutations of x, y, and z. The only difference is a simple but nontrivial factor of two. We can of course eliminate this factor by renormalizing the operators by a factor of $1/2$. The resulting operators are then denoted as the spin operators $\hat{S}$. We define the three components as follows:

$$\begin{aligned} \hat{S}_x &= \frac{1}{2}\hat{O}(\lambda_3); \\ \hat{S}_y &= \frac{1}{2}\hat{O}(\lambda_1); \\ \hat{S}_z &= \frac{1}{2}\hat{O}(\lambda_2). \end{aligned} \tag{8.5}$$

The choice to associate the z-component with $\hat{O}(\lambda_2)$ is based on the fact that this operator acts as the z-component of the angular momentum in the Cartesian $\Sigma(x,y)$-plane. This choice also fixes the remaining $\hat{S}_x$ and $\hat{S}_y$ operators in view of Eq. (8.2). The spin operators have the same commutation relations as the angular momentum operators, and form the generators of the SU(2) group, yielding the commutation relations

$$\left[\hat{S}_k, \hat{S}_l\right] = i\epsilon_{klm}\hat{S}_m. \tag{8.6}$$

The $\hat{S}$ spin operators reconstruct SO(3) in a two-dimensional space.

8.2 SPINORS AND HALF-INTEGER ANGULAR MOMENTUM

The two-dimensional vector space in which the $\hat{S}$ operators act, may be recast in a complex form that diagonalizes the $\hat{S}_z$ operator. These canonical components are said to form a *spinor*. Its components are labeled as α and β. We have:

$$\begin{aligned} a^\dagger_\alpha &= \frac{1}{\sqrt{2}}\left(a^\dagger_x + ia^\dagger_y\right); \\ a^\dagger_\beta &= \frac{1}{\sqrt{2}}\left(a^\dagger_x - ia^\dagger_y\right). \end{aligned} \tag{8.7}$$

The corresponding annihilation operators are then given by

$$\begin{aligned} a_\alpha &= \frac{1}{\sqrt{2}}\left(a_x - ia_y\right); \\ a_\beta &= \frac{1}{\sqrt{2}}\left(a_x + ia_y\right). \end{aligned} \tag{8.8}$$

The internal structure of the electron can be described by such a spinor, which thus gives rise to two states denoted as $|\alpha\rangle$, or spin-up state, versus $|\beta\rangle$, or spin-down state. These states are created from the vacuum by the respective $a^\dagger$ operators.

The spin operators can now be recast in terms of the spin variables by substituting the x and y operators in Eq. (8.1). We can also introduce $\hbar$ as the unit of action, obtaining

$$\begin{aligned}
\hat{S}_x &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\beta + a_\beta^\dagger a_\alpha\right);\\
\hat{S}_y &= \frac{\hbar}{2i}\left(a_\alpha^\dagger a_\beta - a_\beta^\dagger a_\alpha\right);\\
\hat{S}_z &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\alpha - a_\beta^\dagger a_\beta\right).
\end{aligned}\tag{8.9}$$

As an example, the effect of the S_z operator on the spin creation operators is given by

$$\begin{aligned}
\left[\hat{S}_z, a_\alpha^\dagger\right] &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\alpha a_\alpha^\dagger - a_\alpha^\dagger a_\alpha^\dagger a_\alpha\right) = \frac{\hbar}{2} a_\alpha^\dagger;\\
\left[\hat{S}_z, a_\beta^\dagger\right] &= \frac{\hbar}{2}\left(-a_\beta^\dagger a_\beta a_\beta^\dagger + a_\beta^\dagger a_\beta^\dagger a_\beta\right) = -\frac{\hbar}{2} a_\beta^\dagger.
\end{aligned}\tag{8.10}$$

The spin states are thus eigenstates of the $\hat{S}_z$ operator that, in analogy with the $\hat{L}_z$ operator, yields the value of the z-component of angular momentum as

$$\begin{aligned}
\hat{S}_z|\alpha\rangle &= \frac{\hbar}{2}|\alpha\rangle;\\
\hat{S}_z|\beta\rangle &= -\frac{\hbar}{2}|\beta\rangle.
\end{aligned}\tag{8.11}$$

We can thus assign to each spin state a quantum number m_s, which adopts the values $\pm 1/2$. The mathematical treatment in Chapter 4, and even the electronic wavefunction of Möbius molecules in Chapter 5, which led to the conclusion that SO(3) also allows for angular momentum states with half-integer values of m_l, now materializes in the effective spin momenta of the electron. Similarly, we have for the x- and y-components

$$\begin{aligned}
\hat{S}_x|\alpha\rangle &= \frac{\hbar}{2}|\beta\rangle;\\
\hat{S}_x|\beta\rangle &= \frac{\hbar}{2}|\alpha\rangle;
\end{aligned}\tag{8.12}$$

$$\begin{aligned}
\hat{S}_y|\alpha\rangle &= i\frac{\hbar}{2}|\beta\rangle;\\
\hat{S}_y|\beta\rangle &= -i\frac{\hbar}{2}|\alpha\rangle.
\end{aligned}\tag{8.13}$$

These results may be recast in matrix form by using three traceless 2 × 2 matrices known as the *Pauli spin matrices*, which were introduced by Wolfgang Pauli in 1925 in his quantum mechanics of the electron. They are defined as follows:

$$\begin{aligned}
\sigma_z &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix};\\
\sigma_y &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix};\\
\sigma_x &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\end{aligned}\tag{8.14}$$

We may now rewrite the spin operators, acting on the spinor basis (in row vector notation), in terms of the spin matrices:

$$\begin{aligned} \hat{S}_z &= \frac{\hbar}{2}\sigma_z; \\ \hat{S}_y &= \frac{\hbar}{2}\sigma_y; \\ \hat{S}_x &= \frac{\hbar}{2}\sigma_x. \end{aligned} \tag{8.15}$$

Along these lines, we can also produce the total spin operator, which is nothing else than the Casimir operator of SU(2):

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2. \tag{8.16}$$

Application of this operator to, for example, the spin-up state yields

$$\hat{S}^2 \left|\alpha\right\rangle = \frac{3}{4}\hbar^2 \left|\alpha\right\rangle = \frac{1}{2}\left(1+\frac{1}{2}\right)\hbar^2 \left|\alpha\right\rangle. \tag{8.17}$$

The m_s states thus belong to a multiplet with a spin quantum number s, which for the electron equals $1/2$, and characterizes the total spin momentum as $s\,(s+1)\,\hbar^2$, in perfect analogy with the orbital angular momentum states.

The Casimir operator can also be obtained by forming the square norm of the spin matrices, which yields

$$\hat{S}^2 = \frac{\hbar^2}{4}\left(\sigma_x^2+\sigma_y^2+\sigma_z^2\right) = \frac{3}{4}\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{8.18}$$

and confirms the previous result. The 2×2 unit matrix is referred to as σ_0. We will return to these matrices in the second scholium, where they will appear in connection with the quaternion algebra. Extensions of Eq. (8.18) to multiparticle systems are provided in Appendix H.

8.3 THE RELATIONSHIP BETWEEN SU(2) AND SO(3)

Although the $\hat{S}$ and $\hat{L}$ operators give rise to the same Lie algebra, the corresponding groups are not completely identical. In fact, the relationship between the SU(2) and SO(3) groups is a homomorphism instead of an isomorphism, and it is interesting to explore this relationship in more detail because it also helps to clarify the nature of the electron spin. The following treatment is inspired by, for example, the treatment by Griffith,[2] and is based on the interactions in the spinor space.[3] An alternative treatment makes use of the coupling of two spins, as seen, for example, in Altmann.[4]

The interaction Hamiltonian in the space formed by the two spin components is described by a Hermitian matrix—say, $\mathbb{H}$–which, in its most general form, requires three real parameters and can always be written as a linear combination of the

---

[2] J. S. Griffith. *The Theory of Transition-Metal Ions*. Cambridge: Cambridge University Press, 1961.

[3] A. Ceulemans. *Group Theory Applied to Chemistry*. Dordrecht: Springer, 2013.

[4] S. L. Altmann. *Rotations, Quaternions, and Double Groups*. Oxford: Clarendon Press, 1986.

spin matrices. Let us represent the three independent parameters as x, y, and z, corresponding to the labels of the spin matrices:

$$\mathbb{H} = x\sigma_x + y\sigma_y + z\sigma_z = \begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix}. \tag{8.19}$$

In principle, we could add a fourth (time)-component in σ_0, but we have taken $\mathbb{H}$ to be traceless because the trace does not introduce an interaction inside the spin space, but simply shifts the barycenter of the two levels with respect to an external reference. An example of a Hamiltonian is the Zeeman interaction of an isolated spin in a magnetic field. The x, y, and z parameters in this case are proportional to the magnitude of the magnetic field in the three Cartesian directions.

The interaction operator, which gives rise to the Hamiltonian matrix, can be expressed in the spinor basis as

$$\hat{H} = z\left(|\alpha\rangle\langle\alpha| - |\beta\rangle\langle\beta|\right) + x\left(|\alpha\rangle\langle\beta| + |\beta\rangle\langle\alpha|\right) + iy\left(|\beta\rangle\langle\alpha| - |\alpha\rangle\langle\beta|\right). \tag{8.20}$$

This result can also be recast in matrix form as

$$\hat{H} = \begin{bmatrix} |\alpha\rangle & |\beta\rangle \end{bmatrix} \mathbb{H} \begin{bmatrix} \langle\alpha| \\ \langle\beta| \end{bmatrix}, \tag{8.21}$$

or inversely as

$$\mathbb{H} = \begin{bmatrix} \langle\alpha| \\ \langle\beta| \end{bmatrix} \hat{H} \begin{bmatrix} |\alpha\rangle & |\beta\rangle \end{bmatrix}. \tag{8.22}$$

The labeling of the parameters in the Hamiltonian matrix suggests a connection between the two-dimensional complex spin space and a real three-dimensional vector space. To establish this connection between the spinor and the vector, we now need to verify how transformations in the spinor are showing up as transformations in the vector.

Consider a finite element of SU(2), represented by a matrix $\mathbb{U}$. In the special unitary group SU(2), the most general form this matrix can take involves two complex parameters, a and b, subject to the condition that their squared norm, $|a|^2 + |b|^2$, equals unity. These parameters are known as the Cayley-Klein parameters. We get

$$\mathbb{U} = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}. \tag{8.23}$$

The determinant of $\mathbb{U}$ is unity because of the norm condition; hence, $\mathbb{U}$ belongs to the special unitary group SU(2), of all unitary 2 × 2 matrices with determinant +1. The transformation $\hat{U}$ transforms the spinor as follows:

$$\hat{U}\begin{bmatrix} |\alpha\rangle & |\beta\rangle \end{bmatrix} = \begin{bmatrix} |\alpha'\rangle & |\beta'\rangle \end{bmatrix} = \begin{bmatrix} |\alpha\rangle & |\beta\rangle \end{bmatrix}\mathbb{U}. \tag{8.24}$$

To apply the transformation to the interaction operator, we must also consider the effect of $\hat{U}$ on the column of bra functions. This simply requires the inverse of the matrix, which for a unitary matrix is nothing else than its complex conjugate transposed:

$$\hat{U}\begin{bmatrix} \langle\alpha| \\ \langle\beta| \end{bmatrix} = \begin{bmatrix} \langle\alpha'| \\ \langle\beta'| \end{bmatrix} = \mathbb{U}^\dagger \begin{bmatrix} \langle\alpha| \\ \langle\beta| \end{bmatrix}. \tag{8.25}$$

The transformation of the spinor thus changes the interaction matrix as follows:

$$\begin{aligned} \mathbb{H}' &= \begin{bmatrix} \langle\alpha'| \\ \langle\beta'| \end{bmatrix} \hat{H} \begin{bmatrix} |\alpha'\rangle & |\beta'\rangle \end{bmatrix} \\ &= \mathbb{U}^\dagger \begin{bmatrix} \langle\alpha| \\ \langle\beta| \end{bmatrix} \hat{H} \begin{bmatrix} |\alpha\rangle & |\beta\rangle \end{bmatrix} \mathbb{U} \\ &= \mathbb{U}^\dagger \mathbb{H} \mathbb{U}. \end{aligned} \tag{8.26}$$

The transformed Hamiltonian matrix can again be decomposed into the Pauli matrices using a new set of parameters (x', y', z'):

$$\mathbb{H}' = x'\sigma_x + y'\sigma_y + z'\sigma_z. \tag{8.27}$$

Note that the connection between the spinor and vector transformation is made possible by the requirement that the interaction operator itself, $\hat{H}$, is invariant under the transformation. In this way, the transformation of the spinor $(|\alpha\rangle, |\beta\rangle) \rightarrow (|\alpha'\rangle, |\beta'\rangle)$ induces a transformation of the vector $(x, y, z) \rightarrow (x', y', z')$.

In the vector space, this transformation is described by a matrix $\mathbb{O}(U)$. This matrix can be derived by combining the previous two equations. We get

$$(x', y', z') = \hat{O}(U)(x, y, z) = (x, y, z)\,\mathbb{O}(U), \tag{8.28}$$

where the transformation matrix is given by

$$\mathbb{O}(U) = \begin{bmatrix} \frac{1}{2}\left(a^2 + a^{*2} - b^2 - b^{*2}\right) & -\frac{i}{2}\left(a^2 - a^{*2} + b^2 - b^{*2}\right) & -ab - (ab)^* \\ \frac{i}{2}\left(a^2 - a^{*2} - b^2 + b^{*2}\right) & \frac{1}{2}\left(a^2 + a^{*2} + b^2 + b^{*2}\right) & -i\left(ab - (ab)^*\right) \\ ab^* + a^*b & -i(ab^* - a^*b) & |a|^2 - |b|^2 \end{bmatrix}. \tag{8.29}$$

We can now invoke the infinitesimal operators of the SU(2) Lie group to obtain the general expression for the $\mathbb{U}$ matrix. The operator form makes use of the spin operators and reads as follows:

$$\hat{U} = \exp\left(-i\frac{\omega}{\hbar}\,\boldsymbol{n}\cdot\hat{\mathbf{S}}\right). \tag{8.30}$$

The corresponding matrix is given by

$$\mathbb{U} = \exp\left(-i\frac{\omega}{2}\left(n_x\sigma_x + n_y\sigma_y + n_z\sigma_z\right)\right) = \exp\left(-i\frac{\omega}{2}\mathbb{Y}\right). \tag{8.31}$$

Here, $\mathbb{Y}$ is the SU(2) equivalent of the SO(3) $\mathbb{Z}$ matrix:

$$\mathbb{Y} = \begin{bmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{bmatrix}. \tag{8.32}$$

We can then easily verify that

$$\mathbb{Y}^2 = \left(n_x^2 + n_y^2 + n_z^2\right)\mathbb{I} = \mathbb{I}, \tag{8.33}$$

where use was made of the fact that the sum of the squared directional cosines is equal to unity. The matrix exponential thus becomes

$$\begin{aligned} \mathbb{U} &= \cos\omega/2\,\mathbb{I} - i\sin\omega/2\,\mathbb{Y} \\ &= \begin{bmatrix} \cos\omega/2 - in_z\sin\omega/2 & \left(-in_x - n_y\right)\sin\omega/2 \\ \left(-in_x + n_y\right)\sin\omega/2 & \cos\omega/2 + in_z\sin\omega/2 \end{bmatrix}, \end{aligned} \tag{8.34}$$

from which the Cayley-Klein parameters are identified as

$$\begin{aligned} a &= \cos\omega/2 - in_z \sin\omega/2; \\ b &= \left(-in_x - n_y\right)\sin\omega/2. \end{aligned} \tag{8.35}$$

Inserting these results into Eq. (8.29) yields

$$\mathbb{O}(U) = \begin{bmatrix} 1-2\left(n_y^2+n_z^2\right)\sin^2(\omega/2) & -n_z\sin\omega+2n_xn_y\sin^2(\omega/2) & n_y\sin\omega+2n_zn_x\sin^2(\omega/2) \\ n_z\sin\omega+2n_xn_y\sin^2(\omega/2) & 1-2\left(n_z^2+n_x^2\right)\sin^2(\omega/2) & -n_x\sin\omega+2n_yn_z\sin^2(\omega/2) \\ -n_y\sin\omega+2n_zn_x\sin^2(\omega/2) & n_x\sin\omega+2n_yn_z\sin^2(\omega/2) & 1-2\left(n_x^2+n_y^2\right)\sin^2(\omega/2) \end{bmatrix}.$$

This matrix is exactly the general rotation matrix, which was introduced in Chapter 5 (Eq. 5.45), and it describes a rotation over an angle ω about an axis with directional cosines n_x, n_y, and n_z. This results establishes the connection between the unitary transformation of a spinor and the orthogonal transformation of a vector.

The mapping $O(U)$ is not an isomorphism, though, but a homomorphism. Indeed, for the rotations in the spinor space, the angles are halved with respect to the concomitant rotations in the vector space. Hence, although a rotation over 2π in ordinary space corresponds to a return to the original orientation, in spinor space it converts each component into minus itself; the $\mathbb{U}(2\pi)$ matrices are therefore identical to minus the identity matrix, $-\mathbb{I}$, and the mapping between SU(2) and SO(3) is a two-to-one mapping. Each element of the rotation group in three-dimensional space is the image of two elements in SU(2). For this reason, SU(2) is also called the *covering group* of SO(3).

Let us now examine the correspondence between elements of SU(2) and their counterparts in SO(3) for a few canonical examples. First, consider a finite transformation generated by $\hat{S}_z$. Following the method of the Lie exponential, we get

$$\hat{U}_z(\omega) = \exp\left(-i\frac{\omega}{\hbar}\hat{S}_z\right), \tag{8.36}$$

or, in matrix form,

$$\mathbb{U}_z(\theta) = \begin{bmatrix} \exp(-i\omega/2) & 0 \\ 0 & \exp(i\omega/2) \end{bmatrix}. \tag{8.37}$$

Substituting this result in Eq. (8.29) yields a finite transformation in vector space, which we denote as $\mathbb{O}(U_z)$:

$$\mathbb{O}(U_z)\left(x,y,z\right) = \left(x,y,z\right)\begin{bmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{8.38}$$

which is recognized as a rotation around the z-direction over an angle ω. Similar mappings can be found for a rotation around the x-axis:

$$\begin{aligned} \mathbb{U}_x(\omega) &= \begin{bmatrix} \cos\omega/2 & -i\sin\omega/2 \\ -i\sin\omega/2 & \cos\omega/2 \end{bmatrix} \\ \mathbb{O}(U_x) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix}. \end{aligned} \tag{8.39}$$

8.4 THE SPINNING ELECTRON

The special relationship between SU(2) and SO(3) implies that any vector in ordinary space can be considered as a composite entity based on an underlying and more fundamental spinor space. The electron has an intrinsic spinor structure defined by two spin states. When it performs an orbital motion, the intrinsic spin momentum can couple to the angular momentum of the orbital motion based on the common structure of SU(2) and SO(3). As long as the coupling between spin and orbital motion is not included, the symmetry group of the atom remains a formal direct product between whatever dynamic symmetry we apply and SU(2). In the current treatment, we will not be concerned with the fine details of the spin–orbit coupling. Hence, we will not dwell on the internal structure of the electron, but rather concentrate on the grand structure of the atomic shells. Nevertheless, as the next chapter will demonstrate, a role is written for SU(2) in this script as well.

9 The SO(4) group

In One Dimension, did not a moving Point produce a Line with two terminal points? In Two Dimensions, did not a moving Line produce a Square with four terminal points? In Three Dimensions, did not a moving Square produce—did not this eye of mine behold it—that blessed Being, a Cube, with eight terminal points? And in Four Dimensions shall not a moving Cube—alas, for Analogy, and alas for the Progress of Truth, if it be not so—shall not, I say, the motion of a divine Cube result in a still more divine Organization with sixteen terminal points?
–Edwin Abbott Abbott (1884)[1]

On January 17, 1926, the *Zeitschrift für Physik* received a paper from Austrian theoretical physicist Wolfgang Pauli (1900–1958).[2] In his article "On the Hydrogen Spectrum from the Standpoint of the New Quantum Mechanics," Pauli exploited Heisenberg's *matrix mechanics* to derive the spectrum of atomic hydrogen. Besides angular momentum, Pauli also introduced the quantum mechanical analogue of the classical *Laplace-Runge-Lenz vector*, which we will study in §9.3.4. The invariance of the Hamiltonian under these operators proved to be sufficient to explain the complete degeneracy of the hydrogen spectrum. Moreover, the corresponding algebra could be

---

[1] E. A. Abbott. *Flatland: A Romance of Many Dimensions.* London: Seely & Co., 1884.

[2] See W. Pauli. "Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik.", *Zeitschrift für Physik* 36 (1926), pp. 336–363. An English translation of Pauli's paper appeared in W. Pauli. "On the Hydrogen Spectrum from the Standpoint of the New Quantum Mechanics." In: *Sources of Quantum Mechanics.* Ed. B. L. van der Waerden. New York: Dover Publications, 1967, pp. 387–415.

identified as the Lie algebra of the rotational group in four dimensions, isomorphic to the special orthogonal group SO(4). In this chapter, we will reconstruct Pauli's scheme and relate it to the generators of the $\mathfrak{so}(4)$ algebra.[3]

It is amazing to note how Pauli's approach led to the energy eigenvalues of the hydrogen atom, and many of its constraints and features—free from the Schrödinger wave equation, and the accompanying conditions of square integrability. At the time, neither Werner Heisenberg (1901–1976) nor Pauli expected that the whereabouts of the electron in the attractive potential of a proton could be represented by a continuous field, as Erwin Schrödinger's (1887–1961) wave equation was about to show. Schrödinger's first paper on the new *wave mechanics* arrived at the editorial office of the *Annalen der Physik* only ten days later, January 27, 1926.[4]

9.1 THE HYDROGEN ATOM

Since Schrödinger, the standard way to solve the problem of the hydrogen atom has been by brute calculation and by explicit separation of the variables in the governing Schrödinger equation:

$$\hat{\mathscr{H}}\,|nlm\rangle = \left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}\right)|nlm\rangle = E_n\,|nlm\rangle. \tag{9.1}$$

This equation is quite involved, but fortunately, during the preceding era, French mathematicians Pierre-Simon Laplace (1749–1827), Adrien-Marie Legendre (1752–1833), and Edmond Laguerre (1834–1886) had provided all the analytic tools required to solve the eigenvalue problem for the bound states.

9.1.1 A quartet of quantum numbers

The corresponding eigenstates of the Hamiltonian, denoted by the ket $|nlm\rangle$ presented earlier, are uniquely characterized by the quantum numbers n, l, and m_l; the *principal quantum number* n refers to the energy level, and l and m_l are the spherical symmetry labels resulting from the rotational SO(3) group (*vide infra* Chapter 11).[5] The *orbital*

---

[3] Two monographs that deal especially with the symmetries of the hydrogen atom are M. J. Englefield. *Group Theory and the Coulomb Problem.* New York: Wiley-Interscience, 1972; and V. Guillemin and S. Sternberg. *Variations on a Theme by Kepler*, vol. 42. American Mathematical Society Colloquium Publications. Providence: American Mathematical Society, 1990. For a review on group theory and the hydrogen atom, see M. Bander and C. Itzykson. "Group Theory and the Hydrogen Atom (I)." *Reviews of Modern Physics* 38.2 (1966), pp. 330–345 and M. Bander and C. Itzykson. "Group Theory and the Hydrogen Atom (II)." *Reviews of Modern Physics* 38.2 (1966), pp. 346–358. See also C. E. Burkhardt and J. J. Leventhal. *Topic in Atomic Physics.* New York: Springer, 2006; and W. Greiner and B. Müller. *Quantum Mechanics: Symmetries.* Berlin: Springer-Verlag, 2001, pp. 477–496. Other short accounts appear in M. I. Petrashen and E. D. Trifonov. *Applications of Group Theory in Quantum Mechanics.* New York: Dover Publications, 1969, pp. 185–196; A. W. Joshi. *Elements of Group Theory for Physicists.* New Delhi: Wiley Eastern, 1977, pp. 171–176; and H. F. Jones. *Groups, Representations and Physics.* New York: Taylor & Francis, 1998, pp. 124–127. As a mathematical supplement, consider S. Frank Singer. *Linearity, Symmetry, and Prediction in the Hydrogen Atom.* New York: Springer, 2005.

[4] E. Schrödinger. "Quantisierung als Eigenwertproblem." *Annalen der Physik* 386.18 (1926), pp. 109–139.

[5] The principal quantum number n is defined as the sum $n \equiv n_r + l + 1$, with n_r being, the *radial quantum number* (cf. Chapter 11).

angular momentum quantum number l is a measure for the magnitude of the orbital angular momentum L:

$$L = \sqrt{l(l+1)}\hbar. \tag{9.2}$$

In a shell of principal quantum number n, l can take on values from zero to a maximum value that is $n-1$ in integer steps. There are, as a result, n different values of l for a given value of n:

$$\begin{gathered} l = 0,\ 1,\ 2,\ 3,\ \ldots,\ n-1. \\ s,\ p,\ d,\ f,\ \ldots. \end{gathered} \tag{9.3}$$

States with different values of l are distinguished by the lowercase Latin letters s, p, d, and f, as described earlier. A state with $n = 3$ and $l = 1$, for example, is denoted as $3p$. The orbital angular momentum of an electron is complicated further by the fact that it is a *vector* (instead of a *scalar*) and it can therefore point in some direction in space. In quantum mechanics, the *orientation* of the orbital angular momentum vector around an arbitrary axis (e.g., the z-axis) is *quantized* as well. So, for any value of l, the measured value of the angular momentum along the chosen axis (e.g., the z-direction) is given by

$$L_z = m_l \hbar, \tag{9.4}$$

where m_l is called the *magnetic quantum number*, which can take on $2l+1$ different values, ranging from $-l$ to $+l$ in integer steps:

$$m_l = -l,\ -l+1,\ -l+2,\ \ldots,\ l-2,\ l-1,\ l. \tag{9.5}$$

Finally, of quintessential importance, the electron itself has an internal (intrinsic) angular momentum, called *spin*, that is simply a part of the electron. According to quantum mechanics, an electron can have either of two states, spin up or spin down, represented by the arrows ↑ and ↓ or the Greek letters α and β, respectively (cf. Chapter 8). These two spin states are distinguished by a fourth quantum number, the *spin magnetic quantum number* m_s, as first introduced by Wolfgang Pauli (1900–1958). This quantum number can have only two values; $+1/2$ indicates an ↑ electron and $-1/2$ indicates a ↓ electron.

9.1.2 "Accidental" degeneracies and the Fock (*n*) rule

The corresponding energy eigenvalue of the hydrogenic system is quantized with

$$E_n = -\frac{mZ^2e^4}{8h^2{\epsilon_0}^2n^2} \quad \text{for} \quad n = 1,\ 2,\ 3,\ 4,\ \ldots, \tag{9.6}$$

with m the reduced mass, Z the atomic number ($Z = 1$ for hydrogen), e the unit of electric charge, h Planck's constant (the logo of quantum mechanics), and ϵ_0 the vacuum permittivity. It is evident that this spectrum is *manifold degenerate* in l and m_l, dependent only on the quantum number n (Figure 9.1). The total degeneracy (or *multiplicity*) for the principal quantum number n is equal to n^2, according to

$$\sum_{l=0}^{n-1}(2l+1) = (0+(n-1))\,n + \sum_{l=0}^{n-1} 1 = (n-1)\,n + n = n^2. \tag{9.7}$$

Although the Schrödinger equation can be solved exactly for the hydrogen atom, the traditional derivation offers no insight as to the origin of this n^2-fold degeneracy. In

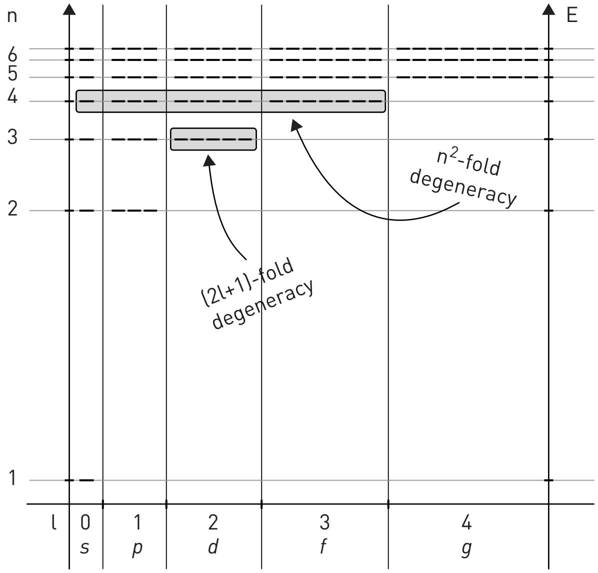


FIGURE 9.1 Energy-level diagram for atomic hydrogen. The $n=1$ state is referred to as the *ground state*, whereas the higher states ($n=2 \rightarrow 7$) are known as *excited states*.

this chapter, we will therefore seek further physical insight into this important system by examining the conditions and conserved properties that lead to this remarkable degeneracy. For the moment, we observe that the energy rule in Eq. (9.6) gives rise to the following orbital sequence for the hydrogen atom:

$$\overbrace{\underbrace{\{1s\}}_{\text{dim}=2}}^{n=1} \ll \overbrace{\underbrace{\{2s=2p\}}_{\text{dim}=8}}^{n=2} \ll \overbrace{\underbrace{\{3s=3p=3d\}}_{\text{dim}=18}}^{n=3} \ll \overbrace{\underbrace{\{4s=4p=4d=4f\}}_{\text{dim}=32}}^{n=4} \ll \ldots. \tag{9.8}$$

We can summarize this sequence of levels of increasing energy by the simple *Fock* (n) *rule.*

Definition 9.1 (The Fock (n) rule): With increasing nuclear charge Z, the nl-orbitals are filled in order of increasing n. ■

9.2 DYNAMICAL SYMMETRIES

The three pillars depicted in Figure 6.6 will form a central theme in this chapter. We already saw that systems with *spherical symmetry* are $(2l+1)$-fold *degenerate* with respect to the z-component of the orbital angular momentum $\hat{L}$. The fact that states of the same l but different m have the same energy is thus accounted for by this symmetry. The degeneracy of the three $2p$-orbitals ($2p_x$, $2p_y$, and $2p_z$ with $l=1$ and $m=-1, 0, 1$), as illustrated in Figure 9.1, is thus a consequence of the hydrogen atom possessing rotational symmetry, as described by the SO(3) group. Since the SO(3) group is generated by the angular momentum operators $\hat{L}_i$ ($i=1, 2, 3$), this symmetry manifests itself to us as the *conservation of angular momentum L*. In brief,

$$\begin{array}{c}\text{Conservation of}\\ \text{Angular Momentum}\end{array} \longleftrightarrow \begin{array}{c}\text{Spherical}\\ \text{Symmetry}\end{array} \longleftrightarrow \begin{array}{c}(2l+1)\text{-fold}\\ \text{Degeneracy}\end{array} \tag{9.9}$$

Interestingly, the hydrogen atom has *additional degeneracy* beyond that associated with rotational symmetry; that is, states of the same n but different l possess the same energy as well. This leads to an n^2-fold degeneracy, rather than the $(2l+1)$-fold degeneracy described above. The $2s$-orbital, for instance, (with $n = 2$ and $l = 0$) belongs to the same energy level as the $2p_x$, $2p_y$, and $2p_z$, orbitals (with $n = 2$ and $l = 1$). Initially, these degeneracies were said to be *accidental* because physicists had no clue as to their exact origin. But, as Harold V. McIntosh recounted in the 1970s: "There [had] always been a feeling that accidental degeneracy might not be so much of an accident after all, in the sense that there might actually have been a *larger group* which would incorporate … the overt [i.e., geometrical] symmetry group."[6]

The idea is thus to look for an additional symmetry of the system, above and beyond the spherical symmetry, that could account for the increased degeneracy of the hydrogen atom. It is clear this symmetry cannot be of a *geometric* nature. Rather, the hydrogen atom seems to be endowed with a *hidden symmetry*, which arises from the particular $1/r$ form of the Coulomb potential. That is to say, although *all* systems with a central potential $V(\mathbf{r}) = V(r)$ are spherically symmetric and thus exhibit the familiar $(2l+1)$-fold degeneracy, the n^2-fold degeneracy is unique for the hydrogen atom, and inherent to the Coulomb potential. This is referred to as an *internal symmetry* or a *dynamical symmetry*.[7] We also expect to find an extra *constant of the motion* associated with this additional dynamical symmetry. In short,

$$\begin{array}{c} \text{Extra Constant} \\ \text{of the Motion} \end{array} \longleftrightarrow \begin{array}{c} \text{Dynamical} \\ \text{Symmetry} \end{array} \longleftrightarrow \begin{array}{c} n^2\text{-fold} \\ \text{Degeneracy} \end{array} \tag{9.10}$$

We are thus faced with the following two interrelated questions: First, what is the nature of this dynamical symmetry that contains the SO(3) group as a subgroup? And second, which observable, other than the angular momentum L, is conserved in time for the hydrogen system?

In the case of geometric symmetries, the nature of a symmetry group can usually be inferred from the corresponding *classical system*. But, for dynamical symmetries, this is not generally the case because most quantum systems have no classical analogue (think of the electron spin in Chapter 8, or the nuclear isospin in Chapter 7). The hydrogen atom, however, represents a special case in that it *does* have a classical analogue under the form of the so-called *Kepler problem*. We thus expect the dynamical symmetries and conserved properties of the hydrogen atom to manifest themselves in a classical formulation of the Kepler problem, and this is why we start this chapter with an in-depth investigation of the hydrogen atom from a classical point of view.

9.3 CLASSICAL KEPLER PROBLEM

We start this section with a brief consideration of *central forces* and *potentials*.

Definition 9.2 (Central force): A force **F** that is radially directed from the center of force O, and whose magnitude F only depends on the distance r from the object to the center,

---

[6] H. V. McIntosh. "Symmetry and Degeneracy." In: *Group Theory and Its Applications*, vol. II. Ed. E. M. Loebl. New York: Academic Press, 1971, p. 79 (emphasis added).

[7] For multielectron systems, the nucleus gets *shielded* as a result of the presence of inner electrons. The effective nuclear Coulomb potential therefore differs from the attractive $1/r$ potential for hydrogen, and the dynamical symmetry gets broken, with the degeneracy lowering to $2l+1$.

is called a *central force*. If we take the origin of a coordinate system as the center of force, and let $\mathbf{r}$ denote the position vector of the object, then

$$\mathbf{F}(\mathbf{r}) = F(r)\,\mathbf{e}_r, \tag{9.11}$$

where $\mathbf{e}_r$ is a unit vector in the radial direction. Moreover, this implies that the force field is *spherically symmetric*. ■

The general relationship between a force $\mathbf{F}$ and the potential energy V in Cartesian coordinates x, y, and z is given by

$$\mathbf{F} \equiv -\nabla V = -\frac{\partial V}{\partial x}\mathbf{e}_x - \frac{\partial V}{\partial y}\mathbf{e}_y - \frac{\partial V}{\partial z}\mathbf{e}_z, \tag{9.12}$$

where ∇ is the *del operator* (or *gradient operator*)[8] and $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are the unit vectors along the three Cartesian directions. Each of the partial derivatives in Eq. (9.12) can be written in terms of the spherical coordinates r, θ, and ϕ by the multivariable chain rule. For example,

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial V}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial V}{\partial \phi}\frac{\partial \phi}{\partial x}. \tag{9.13}$$

However, when dealing with a central force, the corresponding potential is spherically symmetric: $V(\mathbf{r}) = V(r)$. Therefore,

$$\frac{\partial V}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial V}{\partial \phi} = 0, \tag{9.14}$$

which simplifies Eq. (9.13) to

$$\frac{\partial V}{\partial x} = \frac{dV}{dr}\frac{\partial r}{\partial x}. \tag{9.15}$$

Rewriting $\partial r/\partial x$ as

$$\begin{aligned}\frac{\partial r}{\partial x} &= \frac{\partial\left(x^2+y^2+z^2\right)^{1/2}}{\partial x}\\ &= \frac{1}{2}\left(x^2+y^2+z^2\right)^{-1/2} 2x = \frac{x}{r},\end{aligned} \tag{9.16}$$

yields

$$\frac{\partial V}{\partial x} = \frac{x}{r}\frac{dV}{dr}. \tag{9.17}$$

We similarly obtain

$$\frac{\partial V}{\partial y} = \frac{y}{r}\frac{dV}{dr} \quad \text{and} \quad \frac{\partial V}{\partial z} = \frac{z}{r}\frac{dV}{dr}. \tag{9.18}$$

Substituting these results in Eq. (9.12) gives

$$\begin{aligned}\mathbf{F} &= -\frac{x}{r}\frac{dV}{dr}\mathbf{e}_x - \frac{y}{r}\frac{dV}{dr}\mathbf{e}_y - \frac{z}{r}\frac{dV}{dr}\mathbf{e}_z\\ &= -\frac{1}{r}\frac{dV}{dr}\left(x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z\right),\end{aligned} \tag{9.19}$$

or

$$\mathbf{F} = -\frac{dV}{dr}\frac{\mathbf{r}}{r}, \tag{9.20}$$

with $\mathbf{r}/r = \mathbf{e}_r$, the unit vector in the $\mathbf{r}$ direction, confirming that a central force is always radially directed.

---

[8] The squared del operator $\nabla^2 = \nabla\cdot\nabla = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is called the *Laplacian*.

9.3.1 General formulation

In classical Newtonian mechanics, the *Kepler problem* (named after Johannes Kepler, see §9.3.6) is a special case of the two-body problem, where two bodies interact via a *central force* with a strength F that varies as the *inverse square* of the distance r between the two bodies:

$$F(r) = \frac{k}{r^2}, \tag{9.21}$$

where k is a constant proportional to the strength of the force. The Kepler problem arises in numerous situations, but is especially important in celestial mechanics because Newtonian gravity obeys an inverse square force law:

$$F(r) = G\frac{m_1 m_2}{r^2}, \tag{9.22}$$

where G is the gravitational constant, m_1 and m_2 are the masses of the two bodies, and r is the distance between them. Examples include a satellite revolving around a planet (like the moon orbiting Earth), a planet moving about its sun, or binary stars moving about each other.

The Kepler problem is also applicable to the motion of two charged particles, as given by Coulomb's law of electrostatics:

$$F(r) = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}, \tag{9.23}$$

where ϵ_0 is the permittivity of the vacuum, and q_1 and q_2 are the two point charges. The most important example here is the motion of an electron around a proton in the hydrogen atom.

The quantum mechanics of the hydrogen atom, and the celestial mechanics of our solar system, clearly belong to two different realms of reality; they inhabit, so to speak, different (microscopic vs. macroscopic) worlds. Yet, there is a deep connection between both systems in that both are subject to an inverse square force law. We will make great use of this connection. After describing the Kepler problem for the *classical hydrogen atom*, we will then translate our results in quantum mechanical terms and apply them to the *quantum hydrogen atom* and its discrete spectrum of energy states.

The hydrogen atom

The electrostatic interaction between the electron e^- and the hydrogen nucleus p^+ can be described with the aid of Coulomb's law in Eq. (9.23). Since the *scalar* definition of the law merely describes the magnitude of the electrostatic force, the *vector* form is used instead, describing both the magnitude *and* the direction of force **F**. From Eqs. (9.11) and (9.23), we have

$$\mathbf{F}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{(+Ze)(-e)}{r^2}\mathbf{e}_r = -\frac{k}{r^2}\mathbf{e}_r, \tag{9.24}$$

where $\mathbf{e}_r$ is a unit vector along the line connecting the nucleus with the electron (i.e., the force **F** points in the *radial direction*) and

$$k = \frac{Ze^2}{4\pi\epsilon_0} > 0. \tag{9.25}$$

The charge of the electron is denoted by $-e$ whereas the nuclear charge of the hydrogen atom equals $+e$ or, more generally, $+Ze$ for a hydrogen-like ion with a nucleus containing Z protons. The distance between the two charges is given by r. The minus sign in Eq. (9.24) refers to an *attractive force*, providing a potential well centered at the origin; a plus sign, on the other hand, corresponds to a *repulsive force*, such as the force between two charges of equal sign, pushing the interacting particles to infinity. Since we are dealing with the hydrogen atom, let us restrict ourselves to a consideration of the attractive force.

Upon comparison of Eq. (9.20) with Eq. (9.24), we see that

$$\frac{dV}{dr} = \frac{k}{r^2}. \tag{9.26}$$

Integration then yields

$$\int dV = k \int \frac{1}{r^2} dr \tag{9.27}$$

or

$$V = -\frac{k}{r} + C, \tag{9.28}$$

where C is the integration constant. We will put this constant equal to zero so that the zero point of the Coulomb potential V corresponds to the two charges being infinitely separated.

Since $\mathbf{e}_r = \mathbf{r}/r$ (where $\mathbf{r}$ is the position vector of the electron), Eq. (9.24) can be rewritten as follows:

$$\mathbf{F} = -\frac{k}{r^2}\mathbf{e}_r = -\frac{k}{r^3}\mathbf{r}. \tag{9.29}$$

Following Newton's second law, $\mathbf{F} = m\mathbf{a}$, the Coulombic force can also be written as

$$\mathbf{F} = m\mathbf{a} = m\frac{d\mathbf{v}}{dt} = \frac{d\mathbf{p}}{dt}. \tag{9.30}$$

Equating Eq. (9.29) with Eq. (9.30) results in the following differential equation of motion:

$$\frac{d\mathbf{p}}{dt} = -k\frac{\mathbf{r}}{r^3}. \tag{9.31}$$

Whether we study the behavior of planets or the properties of hydrogen, the problem in each of these cases consists of finding the positions and velocities of the two interacting bodies over time. Using classical mechanics, the solution to Eq. (9.31) can then be expressed as a *Kepler orbit.*

A derivation of the equations of motion are deferred to §9.3.5. For now, let us be content with a brief statement of the main results. The solutions to the Kepler problem are *conic sections*: circles, ellipses, parabolas, and hyperbolas. Most familiar among those mathematical curves are the elliptical orbits that the planets in our solar system trace through space (Figure 9.2). The following properties of ellipses can be introduced.

Definition 9.3 (Apoapsis and periapsis): The *apoapsis* or *apocenter* A of an ellipse is defined as the point on the ellipse for which the distance to one of the two foci is greatest; the *periapsis* or *pericenter* P, conversely, represents the point of closest approach to the focus.[9] ■

The straight line AP, connecting the apoapsis with the periapsis, is referred to as the *line of apsides.* This is the *major axis* of the ellipse, which runs through both foci.

---

[9] The terms *perihelion* and *aphelion* can be used when describing the revolution of a planet around the sun, located at one of the foci.

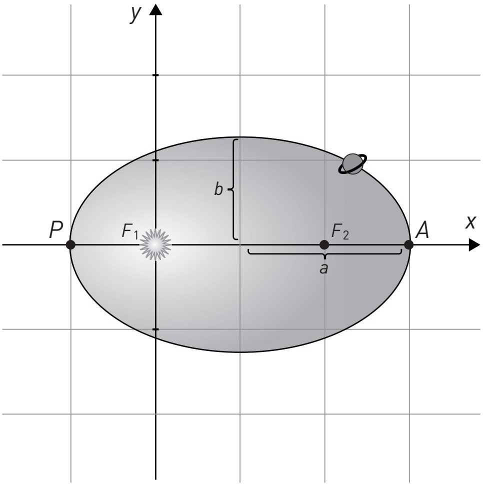


FIGURE 9.2 Elliptical orbit of a planet circling the sun. The sun is located in one of the two foci of the ellipse. The *pericenter* and *apocenter* are denoted P and A, respectively, and the *semimajor* and *semiminor axes* are labeled a and b.

Definition 9.4 (Semimajor and semiminor axes): The *semimajor axis*, denoted a in Figure 9.2, is equal to half the distance from pericenter to apocenter. The *semiminor axis* b is perpendicular to the semimajor axis a. ▪

9.3.2 Hamiltonian formulation

We are now in a position to derive the classical Hamiltonian $\mathscr{H}$ for the hydrogen system. Denoting the masses of the proton and the electron by m_p and m_e, and their velocities by v_p and v_e, respectively, we have

$$\mathscr{H} \equiv T + V = \frac{m_p {v_p}^2}{2} + \frac{m_e {v_e}^2}{2} - \frac{k}{r}. \tag{9.32}$$

The first two terms represent the kinetic energy T of the proton and the electron, whereas the last term V is the contribution of the Coulomb potential, as given in Eq. (9.28).

Reduction to a one-body problem

This two-body system can now be reduced to a *one-body problem* in a central potential; that is, the motion of the two particles can be separated in the uniform motion of the *center of mass* and the relative motion of a single effective particle with respect to the mass center of both. The position of the center of mass is given by

$$\mathbf{R} = \frac{m_p \mathbf{r}_p + m_e \mathbf{r}_e}{m_p + m_e}. \tag{9.33}$$

The velocity of the mass center is thus given by

$$\frac{d\mathbf{R}}{dt} = \frac{m_p \mathbf{v}_p + m_e \mathbf{v}_e}{m_p + m_e}. \tag{9.34}$$

The corresponding kinetic energy T_R is

$$\begin{aligned} T_R = \frac{m_p+m_e}{2}\left(\frac{d\mathbf{R}}{dt}\right)^2 &= \frac{m_p+m_e}{2}\left(\frac{m_p\mathbf{v}_p+m_e\mathbf{v}_e}{m_p+m_e}\right)^2 \\ &= \frac{m_p+m_e}{2}\frac{\left(m_p{}^2v_p{}^2+m_e{}^2v_e{}^2+2m_pm_e\mathbf{v}_p\cdot\mathbf{v}_e\right)}{\left(m_p+m_e\right)^2} \\ &= \frac{m_p{}^2v_p{}^2+m_e{}^2v_e{}^2+2m_pm_e\mathbf{v}_p\cdot\mathbf{v}_e}{2\left(m_p+m_e\right)}. \end{aligned} \tag{9.35}$$

When we subtract the kinetic energy of the mass center from the total kinetic energy T, we obtain the kinetic energy associated with the relative motion:

$$\begin{aligned} T - T_R &= \frac{m_pv_p{}^2}{2}+\frac{m_ev_e{}^2}{2}-\frac{m_p{}^2v_p{}^2+m_e{}^2v_e{}^2+2m_pm_e\mathbf{v}_p\cdot\mathbf{v}_e}{2\left(m_p+m_e\right)} \\ &= \frac{m_pm_ev_p{}^2+m_pm_ev_e{}^2-2m_pm_e\mathbf{v}_p\cdot\mathbf{v}_e}{2\left(m_p+m_e\right)} \\ &= \frac{m_pm_e\left(v_p{}^2+v_e{}^2-2\mathbf{v}_p\cdot\mathbf{v}_e\right)}{2\left(m_p+m_e\right)} \\ &= \frac{1}{2}\frac{m_pm_e}{\left(m_p+m_e\right)}\left(\mathbf{v}_p-\mathbf{v}_e\right)^2. \end{aligned} \tag{9.36}$$

The factor m_pm_e/m_p+m_e is called the *reduced mass*, which we denote for simplicity as m. It represents the mass of an effective particle associated with the relative motion of the electron with respect to the proton. The associated velocity is the relative velocity $\mathbf{v}_p-\mathbf{v}_e$ and is denoted as $\mathbf{v}$. Since the potential energy term only depends on the relative position of the particles with respect to each other, the motion of the center of mass can be separated out effectively, and the Hamiltonian reduces to

$$\mathscr{H} = T+V = \frac{mv^2}{2}-\frac{k}{r} = \frac{p^2}{2m}-\frac{k}{r}. \tag{9.37}$$

9.3.3 Constants of the motion

Our main interest in the Kepler problem concerns the *constants of motion* (i.e., conserved quantities) that will find operator equivalents in the quantum mechanics of the hydrogen problem in §9.4.1.

Conservation of angular momentum L

Evidently, since $\mathscr{H}$ possesses rotational symmetry, the orbital angular momentum $\mathbf{L}=\mathbf{r}\times\mathbf{p}$ is a constant of the motion, as mentioned in Chapter 4:

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}\left(\mathbf{r}\times\mathbf{p}\right) = \mathbf{r}\times\frac{d\mathbf{p}}{dt}+\frac{d\mathbf{r}}{dt}\times\mathbf{p} = 0. \tag{9.38}$$

This results from the fact that $d\mathbf{p}/dt$ is parallel to $\mathbf{r}$ according to Eq. (9.31), and $d\mathbf{r}/dt$ is parallel to $\mathbf{p}$ since $\mathbf{p}=m\mathbf{v}=md\mathbf{r}/dt$. In fact, the invariance of angular momentum implies no less than three conservation laws, because each of the three Cartesian components of this vector is conserved. Notice also that the angular momentum is conserved for *any* central force problem.

Conservation of energy *E*

The total energy E itself, given by $\mathscr{H}$ in equation (9.37), is also independent of time ($d\mathscr{H}/dt = 0$) and therefore represents another constant of the motion. This can be shown by calculating the time derivative of the Hamiltonian:

$$\frac{d\mathscr{H}}{dt} = \frac{d}{dt}\left(\frac{\mathbf{p}^2}{2m} - \frac{k}{r}\right) = \frac{1}{2m}\frac{d\mathbf{p}^2}{dt} - k\frac{d}{dt}\left(\frac{1}{r}\right). \tag{9.39}$$

Using Leibniz's law to find the derivatives of products of functions, as well as the quotient rule to find the derivatives of quotients of functions,

$$\frac{d}{dx}(f.g) = \frac{df}{dx}g + f\frac{dg}{dx}; \tag{9.40}$$

$$\frac{d}{dx}\left(\frac{f}{g}\right) = \frac{\frac{df}{dx}g - f\frac{dg}{dx}}{g^2}. \tag{9.41}$$

Eq. (9.39) can be rewritten as

$$\begin{aligned}\frac{d\mathscr{H}}{dt} &= \frac{1}{2m}\left[\frac{d\mathbf{p}}{dt}\cdot\mathbf{p} + \mathbf{p}\cdot\frac{d\mathbf{p}}{dt}\right] + \frac{k}{r^2}\frac{dr}{dt} \\ &= \frac{\mathbf{p}}{m}\frac{d\mathbf{p}}{dt} + \frac{k}{r^2}\frac{dr}{dt}.\end{aligned} \tag{9.42}$$

To solve the previous equation further, the following equality will prove very useful:

$$\mathbf{r}\cdot\frac{d\mathbf{r}}{dt} = \frac{1}{2}\left(\frac{d\mathbf{r}}{dt}\cdot\mathbf{r} + \mathbf{r}\cdot\frac{d\mathbf{r}}{dt}\right) = \frac{1}{2}\frac{d}{dt}(\mathbf{r}\cdot\mathbf{r}) = \frac{1}{2}\frac{d}{dt}r^2 = \frac{1}{2}\left(\frac{dr}{dt}r + r\frac{dr}{dt}\right) = r\frac{dr}{dt}. \tag{9.43}$$

With the help of Eqs. (9.43) and (9.31), Eq. (9.42) can be rewritten as

$$\begin{aligned}\frac{d\mathscr{H}}{dt} &= \frac{m\mathbf{v}}{m}\left(-k\frac{\mathbf{r}}{r^3}\right) + k\frac{r}{r^3}\frac{dr}{dt} \\ &= -k\frac{\mathbf{r}}{r^3}\frac{d\mathbf{r}}{dt} + k\frac{\mathbf{r}}{r^3}\frac{d\mathbf{r}}{dt} = 0,\end{aligned} \tag{9.44}$$

proving that the energy is indeed a constant of the motion.

9.3.4 The Laplace-Runge-Lenz vector

The conservation of angular momentum, which is a result of the rotational symmetry of the central attraction force, locks the orbit into a plane through the origin O, perpendicular to the direction of $\mathbf{L}$. For convenience, let us take the xy-plane as the orbital plane, with $\mathbf{L}$ pointing in the z-direction.

The Kepler orbits in this plane are subject to additional constraints, however. Indeed, the orbiting particle is confined to an ellipse that is *fixed* in space; that is, the major and minor axes of the ellipse retain a *fixed orientation* in the plane. There is, in other words, no *precession* of the orbit, implying that the ellipse is *closed*. After a full turn, the orbiting particle always returns to its original starting point, retracing its own path *ad infinitum*.

Although the planarity of the orbits is a common feature for *any* central force system, the closure of the Keplerian orbits only occurs for the particular *form* of the central potential for Newtonian attraction, where $V(r) \propto 1/r$. Any departure from

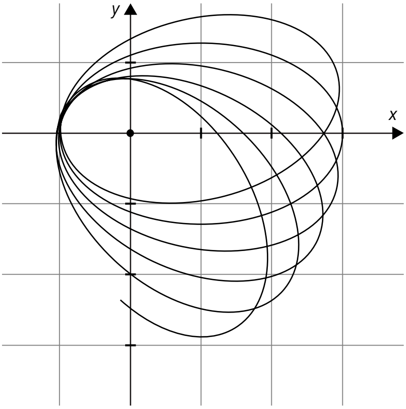


FIGURE 9.3 Precession of the perihelion (I). Any departure from the $1/r$ potential destroys the conservation of the Laplace-Runge-Lenz (LRL) vector. As a consequence, the LRL vector starts to rotate slowly, causing the elliptical orbit to precess.

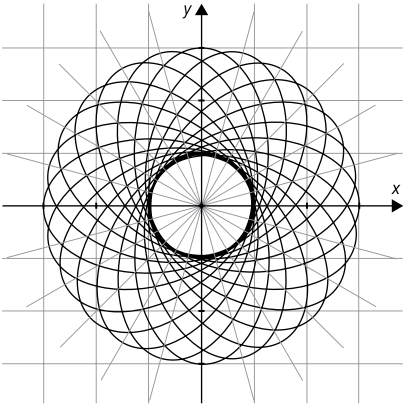


FIGURE 9.4 Precession of the perihelion (II). Over time, a *rosette* is formed, and the orbits are said to be *space-filling*.

this magical potential destroys the closure of the orbits.[10] That is to say, even the smallest deviation of the potential energy from the Newtonian form $V(r) = -k/r$ (to $V(r) = -k/r^{1.1}$, say) would cause the major axis PA of the ellipse to precess slowly. Although the orbit would remain roughly elliptical, it would no longer be closed and a "rosette" would be traced out over time (Figures 9.3 and 9.4). The orbits are then said to be *space-filling*, which is the case for most central forces.

In the words of McIntosh: "The orbits of the Kepler problem are thus almost unique among all the central force problems in that the bounded orbits are simple closed curves."[11] This strongly suggests that there is some additional constant of the motion, other than $\mathscr{H}$ and $\mathbf{L}$, that can be used to characterize the orientation of the major axis in the orbital plane, thus closing the orbit. We expect this to be a constant vector, which we will denote by $\mathbf{M}$ and which should lie along the major axis, pointing from O to P or from O to A.

---

[10] The harmonic oscillator potential $V(r) = 1/2kr^2$ is the only other central potential for which the orbits are closed. This is referred to as *Bertrand's theorem* in honor of the French mathematician Joseph Louis François Bertrand (1822–1900). See also J. Bertrand. "Théorème Relatif au Mouvement d'un Point Attiré vers un Centre Fixe." *Comptes Rendus des Séances de l'Academie des Sciences* 77 (1873), pp. 849–853.

[11] McIntosh, "Symmetry and Degeneracy," p. 81.

Such a vector has indeed been discovered and rediscovered a number of times throughout the history of classical mechanics (see §9.3.7). Among others, the names of Jacob Hermann, Johann Bernoulli, Pierre-Simon Laplace, and William Rowan Hamilton are all associated with this vector. Since Pauli, the name *Laplace-Runge-Lenz vector* (or simply LRL vector) is mostly in use. We write it in the somewhat peculiar form[12]

$$\mathbf{M} \equiv \mathbf{p} \times \mathbf{L} - mk\frac{\mathbf{r}}{r}. \tag{9.45}$$

Conservation of the LRL vector M

The fact that the LRL vector is a constant of the motion is perhaps a little bit more difficult to prove. Taking the time derivative of **M** yields

$$\begin{aligned} \frac{d}{dt}\mathbf{M} &= \frac{d}{dt}\left(\mathbf{p}\times\mathbf{L}\right) - mk\frac{d}{dt}\frac{\mathbf{r}}{r} \\ &= \left[\left(\frac{d\mathbf{p}}{dt}\times\mathbf{L}\right)+\left(\mathbf{p}\times\frac{d\mathbf{L}}{dt}\right)\right] - mk\frac{1}{r^2}\left[\frac{d\mathbf{r}}{dt}r-\mathbf{r}\frac{dr}{dt}\right] \\ &= \mathbf{F}\times\mathbf{L} - \frac{mk}{r^3}\left[\frac{d\mathbf{r}}{dt}r^2 - \mathbf{r}r\frac{dr}{dt}\right], \end{aligned} \tag{9.46}$$

where in the third line, the conservation of angular momentum (Eq. (9.38)) was used, and the fact that $\mathbf{F} = d\mathbf{p}/dt$ (Eq. (9.30)). It will prove interesting to look at the problem for the most general case, where **F** can denote *any* central force (not necessarily of an inverse square nature). Then, by Eq. (9.11),

$$\begin{aligned} \frac{d}{dt}\mathbf{M} &= F(r)\,\frac{\mathbf{r}}{r}\times\mathbf{L} - \frac{mk}{r^3}\left[r^2\frac{d\mathbf{r}}{dt} - \mathbf{r}\left(\mathbf{r}\cdot\frac{d\mathbf{r}}{dt}\right)\right] \\ &= \frac{F(r)}{r}\mathbf{r}\times\left(\mathbf{r}\times\mathbf{p}\right) - \frac{k}{r^3}\left[r^2\mathbf{p} - \mathbf{r}\left(\mathbf{r}\cdot\mathbf{p}\right)\right], \end{aligned} \tag{9.47}$$

where the identity in Eq. (9.43) was invoked in the first line, and $\mathbf{p} = m\mathbf{v} = md\mathbf{r}/dt$ and $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ in the second line. With the help of the vector relation

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}, \tag{9.48}$$

we can rewrite Eq. (9.47) as

$$\begin{aligned} \frac{d}{dt}\mathbf{M} &= \frac{F(r)}{r}\left[(\mathbf{r}\cdot\mathbf{p})\,\mathbf{r} - (\mathbf{r}\cdot\mathbf{r})\,\mathbf{p}\right] - \frac{k}{r^3}\left[r^2\mathbf{p} - \mathbf{r}\left(\mathbf{r}\cdot\mathbf{p}\right)\right] \\ &= \frac{F(r)}{r}\left[(\mathbf{r}\cdot\mathbf{p})\,\mathbf{r} - r^2\mathbf{p}\right] + \frac{k}{r^3}\left[(\mathbf{r}\cdot\mathbf{p})\,\mathbf{r} - r^2\mathbf{p}\right] \\ &= \left[(\mathbf{r}\cdot\mathbf{p})\,\mathbf{r} - r^2\mathbf{p}\right]\left[\frac{F(r)}{r} + \frac{k}{r^3}\right]. \end{aligned} \tag{9.49}$$

It follows that

$$\frac{d}{dt}\mathbf{M} = 0 \quad \Leftrightarrow \quad \frac{F(r)}{r} = -\frac{k}{r^3} \quad \Rightarrow \quad F(r) = -\frac{k}{r^2}. \tag{9.50}$$

That is to say, the LRL vector is a constant of the motion if and only if the force F varies as the *inverse square* of the distance r, which is the case for the Kepler problem, as shown in Eq. (9.24). Actually, because **M** is a vector, it represents no less than three

---

[12] The exact definition of the LRL vector tends to vary from author to author. Some define it as $\mathbf{M} = 1/mk\;\;\mathbf{p}\times\mathbf{L} - \mathbf{r}/r$. Others, such as Pauli, prefer the form $\mathbf{M} = 1/mk\;\;\mathbf{L}\times\mathbf{p} + \mathbf{r}/r$.

conserved quantities through its three Cartesian components M_x, M_y, and M_z. In view of Eq. (9.50), the LRL vector is said to be a *dynamical invariant* because it depends on the particular form of the force law. The angular momentum $\mathbf{L}$ and the energy E, in contrast, are *geometric invariants* because they depend on the geometries of Euclidean space and time.

With our elucidation of the LRL vector as a new constant of the motion, an important piece of the puzzle in Eq. (9.10) falls into place:

$$\begin{array}{c}\text{Conservation of}\\ \text{the LRL vector}\end{array} \longleftrightarrow \begin{array}{c}\text{Dynamical}\\ \text{Symmetry}\end{array} \longleftrightarrow \begin{array}{c}n^2\text{-fold}\\ \text{Degeneracy}\end{array} \tag{9.51}$$

From Noether's theorem, we expect $\mathbf{M}$ to be related to an additional dynamical symmetry of the Kepler problem, beyond the spherical symmetry of all central force problems. The question remains, however: What is the exact nature of this hidden symmetry? We will return to this point in §9.5.

Direction of the LRL vector M

Since $\mathbf{M}$ is conserved for the Kepler problem, its direction in space is fixed. To determine in which direction the LRL vector points, we first note that the vector $\mathbf{p}\times\mathbf{L}$ in Eq. (9.45) is perpendicular to both $\mathbf{p}$ and $\mathbf{L}$. Given that the momentum vector $\mathbf{p}$ is *tangential* to the orbit, and that $\mathbf{L}$ is *perpendicular* to the plane of the orbit, $\mathbf{p}\times\mathbf{L}$ must lie in the xy-plane. Since the unit vector $\mathbf{e}_r = \mathbf{r}/r$ also lies in the plane, $\mathbf{M}$ is forced to lie in the xy-plane.

To determine its position in the xy-plane, we write down the x- and y-components of $\mathbf{M}$ (M_z is, of course, equal to zero). We thus have

$$\begin{aligned}\mathbf{M} &= \mathbf{p}\times\mathbf{L} - mk\frac{\mathbf{r}}{r}\\ &= \begin{vmatrix}\mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z\\ p_x & p_y & p_z\\ L_x & L_y & L_z\end{vmatrix} - \frac{mk}{r}\left(x\mathbf{e}_x + y\mathbf{e}_y\right).\end{aligned} \tag{9.52}$$

This then yields

$$M_x = \left(p_yL_z - p_zL_y\right) - \frac{mk}{r}x; \tag{9.53}$$

$$M_y = \left(p_zL_x - p_xL_z\right) - \frac{mk}{r}y. \tag{9.54}$$

Since the angular momentum vector $\mathbf{L}$ points in the z-direction, $L_x = L_y = 0$. This simplifies the previous equations further to

$$M_x = p_yL_z - \frac{mk}{r}x; \tag{9.55}$$

$$M_y = -p_xL_z - \frac{mk}{r}y. \tag{9.56}$$

Now, let us consider the situation when the orbiting particle (planet or electron) crosses the x-axis. Obviously, $y = 0$ at this point. Also, the momentum vector $\mathbf{p}$ (being tangential to the orbit) points in the y-direction, and thus $p_x = 0$. Substitution in Eq. (9.56) then shows that $M_y = 0$; that is, when crossing the x-axis, the LRL vector $\mathbf{M}$ is pointing in the x-direction. But, because the direction of $\mathbf{M}$ never changes over time, according to Eq. (9.50), this implies that for the elliptical orbit of a classical particle,

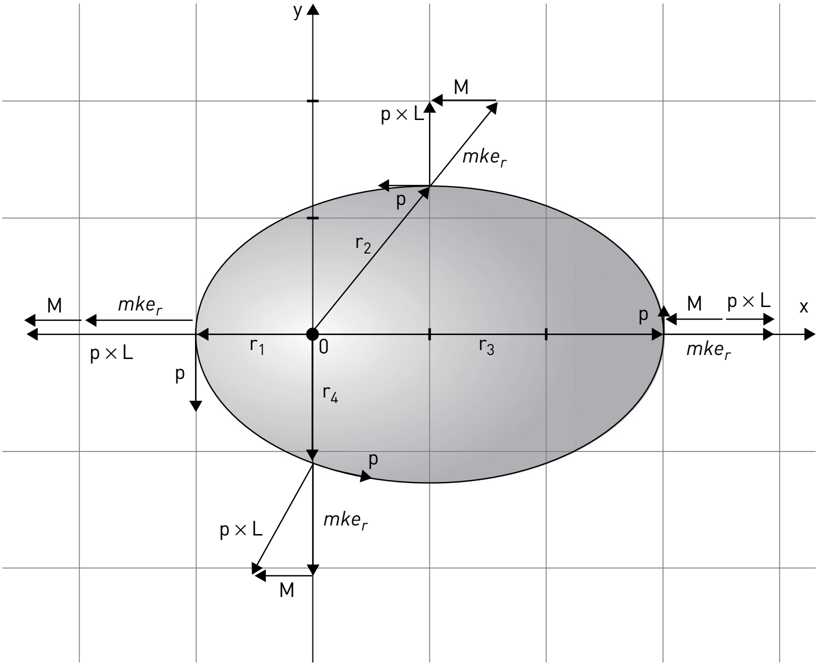


FIGURE 9.5 The orientation of the LRL vector **M** at four different points (denoted by the position vectors $\mathbf{r}_1 \to \mathbf{r}_4$) on a bound elliptical orbit. Since **L** is conserved, **M** lies in the *xy*-plane of a Cartesian coordinate system with the center of attraction as the origin *O*. The coplanar vectors $\mathbf{p} \times \mathbf{L}$ and $mk\mathbf{e}_r$ are indicated for each of the four positions. The momentum vectors **p** are tangential to the orbit. Notice that **M** is constant in both direction and magnitude; that is, **M** is a *constant of the motion* for the Kepler problem.

M is *always* pointing from the focus *O* to the periapsis *P* along the major axis of the ellipse (see Figure 9.5).

Note that for any central force, other than the inverse square force, $d/dt\ \mathbf{M} \neq 0$ according to Eq. (9.50). Since the LRL vector is no longer conserved, it starts rotating in the *xy*-plane and the orbit is seen to precess slowly. This is actually the case in our solar system because the planets are not only attracted to the sun, but they also feel the influence of each other, thus causing a slight deviation of the central potential from the perfect $1/r$ form.[13]

Important relations with the LRL vector

The following two relations will have a direct bearing on the quantum mechanical treatment of the hydrogen atom in §9.4:

$$\mathbf{L} \cdot \mathbf{M} = 0; \tag{9.57}$$

$$\mathbf{M}^2 = 2m\mathscr{H}\mathbf{L}^2 + m^2k^2. \tag{9.58}$$

---

[13] The anomalous precession of the perihelion of the planet Mercury, however, hinted that Newtonian mechanics might not be the entire picture. Although the precession could indeed be explained to a certain extent by the gravitational action of the other planets on Mercury, about 43 arcseconds per century remained unexplained. This opened the doors for Einstein's theory of general relativity, which triumphantly described the precessing perihelion of Mercury.

The orthogonality of **L** and **M** is a direct consequence of their orientation. The vector **L** is defined as the vectorial product $(\mathbf{r}\times\mathbf{p})$ and is thus orthogonal to the orbital plane defined by the radius vector and the displacement of the particle. The vector **M** lies entirely in this plane; its first part, $(\mathbf{p}\times\mathbf{L})$, is orthogonal to **L**; its second part, $-mk\,\mathbf{e}_r$, is along the radius vector. The second proposition in Eq. (9.58) can be proved as follows:

$$\begin{aligned}\mathbf{M}^2 &= \left[\mathbf{p}\times\mathbf{L}-mk\frac{\mathbf{r}}{r}\right]^2\\ &= \left[\mathbf{p}\times(\mathbf{r}\times\mathbf{p})\right]^2-\frac{2mk}{r}\left[\mathbf{p}\times(\mathbf{r}\times\mathbf{p})\right]\cdot\mathbf{r}+m^2k^2\frac{\mathbf{r}\cdot\mathbf{r}}{r^2}.\end{aligned}\tag{9.59}$$

Expanding the triple product $\mathbf{p}\times(\mathbf{r}\times\mathbf{p})$ with the help of Eq. (9.48) yields

$$\mathbf{M}^2=\left[(\mathbf{p}\cdot\mathbf{p})\,\mathbf{r}-(\mathbf{p}\cdot\mathbf{r})\,\mathbf{p}\right]^2-\frac{2mk}{r}\left[(\mathbf{p}\cdot\mathbf{p})\,\mathbf{r}-(\mathbf{p}\cdot\mathbf{r})\,\mathbf{p}\right]\cdot\mathbf{r}+m^2k^2.\tag{9.60}$$

Working out the brackets,

$$\begin{aligned}\mathbf{M}^2 &= \left[p^4r^2-2\,(\mathbf{p}\cdot\mathbf{r})^2\,p^2+(\mathbf{p}\cdot\mathbf{r})^2\,p^2\right]-\frac{2mk}{r}\left[p^2r^2-(\mathbf{p}\cdot\mathbf{r})^2\right]+m^2k^2\\ &= 2m\left[\frac{p^2}{2m}-\frac{k}{r}\right]\left[p^2r^2-(\mathbf{p}\cdot\mathbf{r})^2\right]+m^2k^2.\end{aligned}\tag{9.61}$$

After introducing the Hamiltonian (Eq. (9.37)) and the square of the angular momentum $\mathbf{L}^2=p^2r^2-(\mathbf{p}\cdot\mathbf{r})^2$ (as derived in Eq. (5.135)), we finally obtain

$$\mathbf{M}^2=2m\mathscr{H}\mathbf{L}^2+m^2k^2.\tag{9.62}$$

9.3.5 Equations of motion

In what follows, the classical equations of motion will be derived. With the definition of the LRL vector in Eq. (9.45), this becomes a straightforward matter. Taking the scalar product of the LRL vector **M** with the position vector **r** yields

$$\begin{aligned}\mathbf{M}\cdot\mathbf{r} &= \left(\mathbf{p}\times\mathbf{L}-mk\frac{\mathbf{r}}{r}\right)\cdot\mathbf{r}\\ &= (\mathbf{p}\times\mathbf{L})\cdot\mathbf{r}-mk\frac{\mathbf{r}\cdot\mathbf{r}}{r}\\ &= (\mathbf{r}\times\mathbf{p})\cdot\mathbf{L}-mk\frac{r^2}{r}\\ &= \mathbf{L}\cdot\mathbf{L}-mkr.\end{aligned}\tag{9.63}$$

In the third line of Eq. (9.63), we used the following vector identity:

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=(\mathbf{c}\times\mathbf{a})\cdot\mathbf{b}=(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}.\tag{9.64}$$

We can also write

$$\mathbf{M}\cdot\mathbf{r}=Mr\cos\theta,\tag{9.65}$$

where $M=|\mathbf{M}|$ is the magnitude of the LRL vector and θ is the angle between the two vectors **M** and **r**. Equating Eq. (9.63) with Eq. (9.65), we obtain

$$\begin{aligned}L^2 &= Mr\cos\theta+mkr\\ &= (M\cos\theta+mk)\,r,\end{aligned}\tag{9.66}$$

or

$$r = \frac{L^2}{M\cos\theta + mk}. \tag{9.67}$$

Dividing the numerator and the denumerator by mk yields

$$r = \frac{\frac{L^2}{mk}}{\frac{M}{mk}\cos\theta + 1}. \tag{9.68}$$

After defining

$$\frac{L^2}{mk} \equiv \alpha \quad \text{and} \quad \frac{M}{mk} \equiv \varepsilon, \tag{9.69}$$

the previous equation reduces to

$$r = \frac{\alpha}{1 + \varepsilon\cos\theta}. \tag{9.70}$$

Quite remarkably, this turns out to be the formula for a *conic section* in polar coordinates, with the origin O at one of the foci. In particular, when Eq. (9.70) is applied to the planets in our solar system, the variables r and θ form the polar coordinates (r,θ) of a vector that points from the sun (in one of the foci) to the planet, and traces an ellipse, as shown in Figure 9.6. We have, in other words, obtained a most elegant and simple proof of Kepler's first law of planetary motion.

Definition 9.5 (Kepler's first law): The orbit of every planet is an *ellipse* with the sun at one of the two foci. ■

The parameter α in Eq. (9.70) is called the *semilatus rectum*, whereas ε is known as the *orbital eccentricity* of the ellipse. The eccentricity is a measure for the "ellipticity" of the orbit; it tells you, in other words, how much the planetary orbit deviates from a perfect circle. An orbit is entirely circular when the eccentricity equals zero: $\varepsilon = 0$. This can be easily seen by expressing the variables r and θ in terms of the Cartesian coordinates x and y. From Eq. (9.70), we have

$$\alpha = r + \varepsilon r\cos\theta. \tag{9.71}$$

Given that $r\cos\theta = x$, we get

$$r = \alpha - \varepsilon x. \tag{9.72}$$

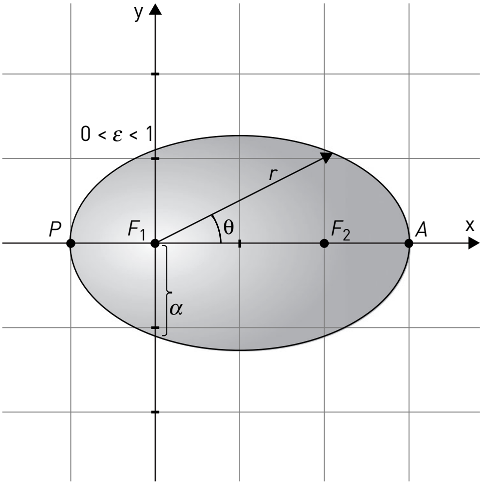


FIGURE 9.6 Ellipse in polar coordinates. Each point on the ellipse is defined by the length of the position vector **r** and the angle θ. For an ellipse the *orbital eccentricity* ε obeys the relation $0 < \varepsilon < 1$. The *semilatus rectum* is denoted α.

Squaring yields

$$r^2 = \alpha^2 - 2\varepsilon\alpha x + \varepsilon^2 x^2. \tag{9.73}$$

By Pythagoras, $r^2 = x^2 + y^2$, yielding the equation for the orbit in Cartesian coordinates:

$$\left(1-\varepsilon^2\right)x^2 + 2\varepsilon\alpha x + y^2 = \alpha^2. \tag{9.74}$$

Setting $\varepsilon = 0$, the previous equation reduces to

$$x^2 + y^2 = \alpha^2, \tag{9.75}$$

which is indeed the equation of a circle with radius α, centered at the origin. Usually, however, $0 < \varepsilon < 1$, and the orbit becomes elliptical; that is, as ε tends toward 1, the orbit gets more and more elongated.[14] To see this, we can multiply each term in Eq. (9.74) by $1-\varepsilon^2$:

$$\left(1-\varepsilon^2\right)^2 x^2 + 2\varepsilon\alpha\left(1-\varepsilon^2\right)x + \left(1-\varepsilon^2\right)y^2 = \alpha^2\left(1-\varepsilon^2\right), \tag{9.76}$$

and add the term $\varepsilon^2\alpha^2$ on both sides of the equation:

$$\left(1-\varepsilon^2\right)^2 x^2 + 2\varepsilon\alpha\left(1-\varepsilon^2\right)x + \varepsilon^2\alpha^2 + \left(1-\varepsilon^2\right)y^2 = \alpha^2\left(1-\varepsilon^2\right) + \varepsilon^2\alpha^2. \tag{9.77}$$

When divided by $\left(1-\varepsilon^2\right)^2$, we obtain

$$x^2 + 2\frac{\varepsilon\alpha}{\left(1-\varepsilon^2\right)}x + \frac{\varepsilon^2\alpha^2}{\left(1-\varepsilon^2\right)^2} + \frac{1}{\left(1-\varepsilon^2\right)}y^2 = \frac{\alpha^2}{\left(1-\varepsilon^2\right)^2}. \tag{9.78}$$

This can be rewritten as

$$\left[x + \frac{\varepsilon\alpha}{\left(1-\varepsilon^2\right)}\right]^2 + \frac{1}{\left(1-\varepsilon^2\right)}y^2 = \frac{\alpha^2}{\left(1-\varepsilon^2\right)^2}, \tag{9.79}$$

or

$$\frac{\left[x + \frac{\varepsilon\alpha}{\left(1-\varepsilon^2\right)}\right]^2}{\frac{\alpha^2}{\left(1-\varepsilon^2\right)^2}} + \frac{y^2}{\frac{\alpha^2}{\left(1-\varepsilon^2\right)}} = 1. \tag{9.80}$$

The expression in Eq. (9.80) corresponds to the equation for an ellipse in Cartesian coordinates:

$$\frac{(x-a)^2}{c^2} + \frac{\left(y-b\right)^2}{d^2} = 1, \tag{9.81}$$

where a, b, c and d are parameters. It represents, in other words, the Cartesian equivalent of Eq. (9.70) for $0 < \varepsilon < 1$, and describes an ellipse centered at the point $\left(-\varepsilon\alpha/\left(1-\varepsilon^2\right), 0\right)$. The *minor radius* along the x-axis has a length of $\alpha/\left(1-\varepsilon^2\right)$, whereas the *major radius* along the y-axis equals $\alpha/\sqrt{1-\varepsilon^2}$.

In a similar vein, it can be shown that Eq. (9.74) describes a *parabola* for $\varepsilon = 1$, and a *hyperbola* for $\varepsilon > 1$.[15] Furthermore, there is an interesting relationship

---

[14] The eccentricities of the planets in our solar system are as follows: $\varepsilon_{\text{Mercury}} = 0.20563069$, $\varepsilon_{\text{Venus}} = 0.00677323$, $\varepsilon_{\text{Earth}} = 0.01671022$, $\varepsilon_{\text{Mars}} = 0.09341233$, $\varepsilon_{\text{Jupiter}} = 0.04839266$, $\varepsilon_{\text{Saturn}} = 0.05415060$, $\varepsilon_{\text{Uranus}} = 0.04716771$, and $\varepsilon_{\text{Neptune}} = 0.00858587$. With the exception of $\varepsilon_{\text{Mercury}}$, all eccentricities are close to zero, accounting for the near circularity of most planetary orbits. Many asteroids, on the other hand, have eccentricities between 0 and 0.30, and Halley's comet has an eccentricity of 0.967.

[15] Some comets entering our solar system, for instance, have been observed to follow near-parabolic orbits, or even hyperbolic paths. When $\varepsilon < 1$, the comets trace an ellipse and return periodically; they are called *periodic comets*. Comets with $\varepsilon \geq 1$, however, are referred to as *nonperiodic* comets.

Table 9.1 Dependency of eccentricity ε on total energy E of the system

Energy	*Eccentricity*	*Conic Section*
$E < 0$	$\varepsilon < 1$	Ellipse
$E = 0$	$\varepsilon = 1$	Parabola
$E > 0$	$\varepsilon > 1$	Hyperbola

between the eccentricity ε and the energy E of the system. To this end, we rewrite Eq. (9.58) as

$$M^2 = 2mEL^2 + m^2k^2, \tag{9.82}$$

or on dividing by m^2k^2,

$$\left(\frac{M}{mk}\right)^2 = 2\left(\frac{L^2}{mk^2}\right)E + 1. \tag{9.83}$$

With the help of the defining relation for eccentricity ε in Eq. (9.69), we obtain

$$\varepsilon^2 = 2\beta E + 1, \tag{9.84}$$

where $\beta = L^2/mk^2 \geq 0$. The dependency of ε on the total energy E of the system is summarized in Table 9.1. It follows that closed elliptical orbits occur only for *bound states* where $E < 0$. When, in contrast, $E > 0$, the object in question traces a hyperbola and we see a *scattering state*. For a hydrogen atom in its *ground state*, for instance, the electron is bound to the proton and the energy $E < 0$ ($E = -13.6$ eV, to be exact). We also say that the electron is *trapped* inside the *potential well* formed by the electromagnetic field that surrounds the positively charged nucleus. By pumping energy into the system, the electron gets excited to higher energies, tracing increasingly elongated orbits. As soon as $E > 0$, the hydrogen atom loses its electron forever and gets ionized ($\mathrm{H} \rightarrow \mathrm{H}^+ + e^-$).

We end the classical description of the hydrogen atom by noting that the magnitude M of the LRL vector $\mathbf{M}$ is proportional to the eccentricity of the orbit, following Eq. (9.69):

$$M = \varepsilon mk. \tag{9.85}$$

For this reason, the LRL vector can be rescaled to yield the so-called *eccentricity vector*:

$$\boldsymbol{\varepsilon} \equiv \frac{\mathbf{M}}{mk} = \frac{1}{mk}\mathbf{p} \times \mathbf{L} - \frac{\mathbf{r}}{r}. \tag{9.86}$$

9.3.6 History of the Kepler problem

The Kepler problem is named after the German mathematician and astronomer Johannes Kepler (1571–1630). While a student at the University of Tübingen, Kepler became an ardent advocate of the heliocentric ideas of Nicolaus Copernicus (1473–1543). In his first major work on astronomy, the *Mysterium Cosmographicum* (published in 1596),[16] Kepler presented a model of the solar system with the sun at the center. Like so many of the notable figures of the seventeenth-century scientific

---

[16] J. Kepler. *Mysterium Cosmographicum.* Tübingen, 1596.

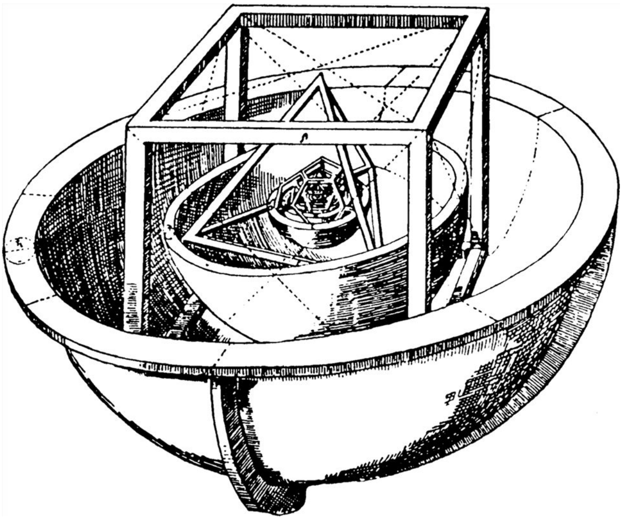


FIGURE 9.7 Kepler's Platonic solid model of the solar system from his *Mysterium Cosmographicum* (1596).

revolution, Kepler, too, was deeply influenced by the musings of the great philosophers of Classical Antiquity. Impressed by the supreme beauty and perfection of the Platonic solids, Kepler attempted to nest these polyhedra into one another like a Russian matryoshka doll (Figure 9.7). By inscribing and circumscribing the five Platonic solids by spherical orbits, Kepler obtained six layers that (quite miraculously) corresponded to the planetary orbits of the planets Mercury, Venus, Earth, Mars, Jupiter, and Saturn. It should not come as a surprise that Kepler assumed the planetary orbits to be perfectly circular. After all, the circle represented the Platonic ideal of symmetry and perfection.

Kepler's work soon attracted the attention of Danish astronomer and alchemist with brass nose, Tycho Brahe (1546–1601), who was known far and wide for his impressive accumulation of astronomical data during the last few decades of his life. Although a disbeliever of the Copernican system, Brahe was deeply impressed by the mathematical skills of Kepler, and invited him to come and work at his research institute Uraniborg on the island of Hven. Brahe hoped Kepler would prove his own Tychonic system, but Kepler decided otherwise. When Tycho unexpectedly died in 1601, Kepler took hold of Tycho's empirical data and started analyzing Tycho's observations of the planet Mars, which had plagued astronomers for centuries.

Kepler's meticulous calculations were deeply troublesome, however. Try as he might, Kepler was unable to fit Mars' orbit in a perfect circle. In 1605, Kepler finally stumbled on the idea that the planetary orbits might not be circles after all, but *ellipses*, with the sun in one of the two foci! His first two laws of planetary motion were published in 1609 in his *Astronomia Nova*[17] and marked the end of Kepler's cherished belief in the perfect circularity of the planetary orbits. Giving up this idea—grounded in the crystalline symmetries of Greek philosophy—did not come easily. It took Kepler eight entire years of work with hundreds of pages of painstaking calculations before he was convinced of the truth of his newfangled idea. And yet, Kepler's initial belief

---

[17] J. Kepler. *Astronomia Nova.* Heidelberg: G. Voegelinus, 1609.

in perfect circles was not completely unfounded. Nearly 250 years later, Sir William Rowan Hamilton (1805–1865) found a way to regain the circularity of the planetary orbits. We will come back to this intriguing insight in §9.6.3.

9.3.7 History of the LRL vector (I)

It is somewhat of an irony that the LRL vector was named after Laplace, Runge, and Lenz, given that none of these scholars actually discovered it. The LRL vector never became as well established in scientific circles as, say, the more familiar momentum or angular momentum vectors, which probably explains why it was discovered and rediscovered a number of times throughout the centuries.[18]

The inverse Newton problem

In any case, it seems that the LRL vector made its first appearance in summer 1710 in a letter that Jakob Hermann (1678–1733) wrote to the eminent Swiss mathematician Johann Bernoulli (1667–1748). Both Hermann and Bernoulli were familiar with Newton's demonstration of the inverse square force law of gravitational attraction from Kepler's three laws of planetary motion—an extraordinary tour de force that came to be known as "*le problème direct*" (or the direct problem). But, Hermann set forth to solve "*le problème inverse*" which he considered even more challenging and mind-bending than the direct problem. The aim, then, was to derive Kepler's laws from Newton's force law. To be more specific, Hermann attempted to prove that conic sections were the only solutions to the planetary orbits under the $1/r^2$ attraction of the sun. In brief, the situation looked as follows:

$$\text{Newton's Force Law} \xrightleftharpoons[\text{direct}]{\text{inverse}} \text{Kepler's Laws}$$

In his derivation, there were signs of what was later recognized as the LRL vector, and Hermann showed it was a constant of the motion for the Kepler problem. His letter, dated July 12, 1710, was sent from Padua to Basel, where Bernoulli spent his later years as a professor of mathematics. Bernoulli, in turn, generalized the results obtained by Hermann and derived the direction and magnitude of the LRL vector. Both Hermann's letter and Bernoulli's answer (dated October 7, 1710) were later sent to the Paris Academy for publication in the *Histoire de l'Académie Royale des Sciences*.[19]

Laplace's *Traité de Mécanique Céleste*

Toward the end of the century, the LRL vector surfaced again in a monumental five-volume exposition on celestial mechanics: the *Traité de Mécanique Céleste* by Pierre-Simon de Laplace (1749–1827).[20] Throughout his work, Laplace applied the

---

[18] For a detailed account, see H. Goldstein. "Prehistory of the Runge-Lenz Vector." *American Journal of Physics* 43.8 (1975), pp. 737–738 and H. Goldstein. "More on the Prehistory of the Laplace or Runge-Lenz Vector." *American Journal of Physics* 44.11 (1976), pp. 1123–1124.

[19] See J. Hermann. "Extrait d'Une Lettre de M. Herman à M. Bernoulli, Datée de Padoüe le 12. Juillet 1710." *Histoire de l'Académie Royale des Sciences* (1732), pp. 519–521 and J. Bernoulli. "Extrait de la Réponse de M. Bernoulli à M. Herman, Datée de Basle le 7. Octobre 1710." *Histoire de l'Académie Royale des Sciences* (1732), pp. 521–533.

[20] P. S. Laplace. *Traité de Mécanique Celeste*, vol. I. L. Paris: J. B. M. Duprat, 1799.

tools of differential calculus, approaching the solar system in an analytic, rather than a geometric, way as Newton had done in his *Principia*.[21] In the first volume, Laplace determined the seven constants of the motion for the Kepler problem—**L**, **M** and E—and used them to derive Kepler's laws in much the same way as we did in §9.3.5.[22]

About fifty years later, in July 1845, William Rowan Hamilton (1805–1865) derived the eccentricity vector in an article titled "Applications of Quaternions to Some Dynamical Questions".[23] In a closely related article, he also introduced the concept of a *hodograph*, which will be discussed in §9.6.3.[24]

The first formulation of the LRL vector in its modern vector notation had to await the *Vector Analysis*, by Josiah Willard Gibbs (1839–1903), published in 1901.[25] About twenty years later, Carl Runge published his popular monograph *Vektoranalysis*, in which he offered a derivation of the Kepler orbits in Laplace style.[26] This formed the main source for Lenz's and Pauli's use of the LRL vector in their quantum mechanical study of the hydrogen-atom (see §9.6.2).

In conclusion, it seems to be more correct to use the name Hermann-Bernoulli-Laplace-Hamilton-Runge-Lenz vector (as one author actually did[27]). Nevertheless, let us opt for the more familiar term Laplace-Runge-Lenz vector.

9.4 QUANTUM MECHANICS OF THE HYDROGEN ATOM

We have seen that the classical Kepler problem carries *seven invariants*: besides the energy $\mathscr{H}$ and the three components L_x, L_y, and L_z of the angular momentum vector **L** (representing the *geometric invariants* for *any* central problem), the Kepler/Coulomb problem is characterized by three additional conserved quantities (i.e., the *dynamical invariants*), which we identified as the components M_x, M_y, and M_z of the LRL vector **M**. In 1926, Pauli constructed the corresponding quantum mechanical equivalents of these constants of the motion, and he subsequently derived the hydrogen spectrum from their commutation algebra, as we shall see in this and the following sections.

---

[21] I. Newton. *The Principia, Mathematical Principles of Natural Philosophy: A New Translation by I. B. Cohen and Anne Whitman*. Berkeley: University of California Press, 1999.

[22] Referred to as the *French Newton*, Laplace's genius soon attracted the attention of Napoleon Bonaparte, who met him at a reception in Josephine Bonaparte's rose garden. "Someone had told Napoleon that the book contained no mention of the name of God; Napoleon, who was fond of putting embarrassing questions, received it with the remark, 'M. Laplace, they tell me you have written this large book on the system of the universe, and have never even mentioned its Creator.' Laplace, who, though the most supple of politicians, was as stiff as a martyr on every point of his philosophy, drew himself up and answered bluntly, 'Je n'avais pas besoin de cette hypothèse-là.' ('I had no need of that hypothesis.') Napoleon, greatly amused, told this reply to Lagrange, who exclaimed, 'Ah! c'est une belle hypothèse; ça explique beaucoup de choses.' ('Ah! It is a fine hypothesis; it explains many things.') Quoted from W. W. Rouse Ball. *A Short Account of the History of Mathematics*. New York: Dover Publications, 1960, p. 343.

[23] W. R. Hamilton. "Applications of Quaternions to Some Dynamical Questions." *Proceedings of the Royal Irish Academy* 3 (1847), pp. xxxvi–l.

[24] W. R. Hamilton. "The Hodograph or a New Method of Expressing in Symbolic Language the Newtonian Law of Attraction." *Proceedings of the Royal Irish Academy* 3 (1847), pp. 344–353.

[25] J. W. Gibbs and E. B. Wilson. *Vector Analysis*. New York: Scribners and Sons, 1901, p. 135.

[26] C. Runge. *Vektoranalysis*, vol. 1. Leipzig: Verlag S. Hirzel, 1919.

[27] See P. R. Subramanian. "Hermann-Bernoulli-Laplace-Hamilton-Runge-Lenz Vector." *Physics Education* 7.4 (1991), pp. 323–327.

9.4.1 Operators for conserved quantities

To treat the hydrogen atom, the foregoing quantities **r**, **p**, **L**, and **M** must all be translated into quantum mechanical terms by converting these vectors into Hermitian operators. This has already been done before for **r** and **p**, where $\hat{\mathbf{r}} = \mathbf{r}$ is the position operator, and $\hat{\mathbf{p}}$ is the momentum operator, given by

$$\hat{\mathbf{p}} \equiv -i\hbar\nabla = \frac{\hbar}{i}\left(\frac{\partial}{\partial x}\mathbf{e}_x + \frac{\partial}{\partial y}\mathbf{e}_y + \frac{\partial}{\partial z}\mathbf{e}_z\right). \tag{9.87}$$

Position and momentum are conjugate operators and they obey *Heisenberg's canonical commutation relations*:

$$\left[\hat{r}_i, \hat{p}_j\right] = i\hbar\delta_{ij}, \tag{9.88}$$

where δ_{ij} is the Kronecker delta, as derived in Appendix I.

Naturally, $\mathscr{H}$, **L**, and **M** will become operators as well within the quantum mechanical description of the hydrogen atom. We obtain $\hat{\mathscr{H}}$ by using Eq. (9.37):

$$\hat{\mathscr{H}} = \frac{\hat{p}^2}{2m} - \frac{k}{r} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{k}{r}. \tag{9.89}$$

The components of the angular momentum operator were defined previously in Chapter 4. When attempting to find a quantum analogue for the classical LRL vector **M**, however, a new problem arises since $\hat{\mathbf{L}}$ and $\hat{\mathbf{p}}$ do *not* commute (see Appendix I):

$$\left[\hat{L}_i, \hat{p}_j\right] = i\hbar\varepsilon_{ijk}\hat{p}_k. \tag{9.90}$$

In view of this, it is not immediately clear in which order the operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{p}}$ have to be placed in the definition of the LRL operator $\hat{\mathbf{M}}$ since $\hat{\mathbf{L}} \times \hat{\mathbf{p}}$ is different from $-\hat{\mathbf{p}} \times \hat{\mathbf{L}}$. If we simply take the operator equivalent of the classical definition

$$\hat{\mathbf{M}} = \hat{\mathbf{p}} \times \hat{\mathbf{L}} - mk\frac{\hat{\mathbf{r}}}{r}, \tag{9.91}$$

the operator $\hat{\mathbf{M}}$ is not Hermitian; that is, $\hat{\mathbf{M}} \neq \hat{\mathbf{M}}^\dagger$. Neither is the alternative operator

$$\hat{\mathbf{M}} = -\hat{\mathbf{L}} \times \hat{\mathbf{p}} - mk\frac{\hat{\mathbf{r}}}{r}. \tag{9.92}$$

Pauli solved this "ordering ambiguity" by *symmetrizing* the previous expressions to ensure Hermiticity.[28] In this way, the classical equation (9.45) debouches into the following definition of the quantum mechanical LRL operator $\hat{\mathbf{M}}$:

$$\hat{\mathbf{M}} = \frac{1}{2}\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - \hat{\mathbf{L}} \times \hat{\mathbf{p}}\right) - mk\frac{\hat{\mathbf{r}}}{r}. \tag{9.93}$$

It is clear that this definition transforms into the classical LRL vector (Eq. (9.91)) if the operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{p}}$ commute. The operator $\hat{\mathbf{M}}$ also reduces to Eq. (9.91) in the classical limit as $\hbar \rightarrow 0$.

---

[28] Jones, *Groups, Representations and Physics*, p. 125.

The Hermiticity of the LRL operator $\hat{\mathbf{M}}$

The Hermiticity of $\hat{\mathbf{M}}$, as defined in Eq. (9.93), is as follows. First, let us consider the x-component of $\hat{\mathbf{M}}$:

$$\begin{aligned}\hat{M}_x &= \frac{1}{2}\left[\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_x - \left(\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)_x\right] - mk\frac{x}{r}\\ &= \frac{1}{2}\left[\left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y\right) - \left(\hat{L}_y\hat{p}_z - \hat{L}_z\hat{p}_y\right)\right] - mk\frac{x}{r}.\end{aligned} \tag{9.94}$$

The Hermitian conjugate of $\hat{M}_x$ then equals

$$\begin{aligned}\hat{M}_x^\dagger &= \frac{1}{2}\left[\left(\hat{p}_y\hat{L}_z\right)^\dagger - \left(\hat{p}_z\hat{L}_y\right)^\dagger - \left(\hat{L}_y\hat{p}_z\right)^\dagger + \left(\hat{L}_z\hat{p}_y\right)^\dagger\right] - mk\frac{x^\dagger}{r}\\ &= \frac{1}{2}\left[\hat{L}_z^\dagger\hat{p}_y^\dagger - \hat{L}_y^\dagger\hat{p}_z^\dagger - \hat{p}_z^\dagger\hat{L}_y^\dagger + \hat{p}_y^\dagger\hat{L}_z^\dagger\right] - mk\frac{x^\dagger}{r}.\end{aligned} \tag{9.95}$$

Since the position, momentum, and angular momentum operators are all Hermitian, we obtain

$$\begin{aligned}\hat{M}_x^\dagger &= \frac{1}{2}\left[\hat{L}_z\hat{p}_y - \hat{L}_y\hat{p}_z - \hat{p}_z\hat{L}_y + \hat{p}_y\hat{L}_z\right] - mk\frac{x}{r}\\ &= \frac{1}{2}\left[\left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y\right) - \left(\hat{L}_y\hat{p}_z - \hat{L}_z\hat{p}_y\right)\right] - mk\frac{x}{r} = \hat{M}_x.\end{aligned} \tag{9.96}$$

The Hermiticity of $\hat{M}_y$ and $\hat{M}_z$ can be demonstrated analogously.

Relations between $\hat{\mathbf{L}}$, $\hat{\mathbf{M}}$, and $\hat{\mathscr{H}}$

Pauli furthermore established the quantum analogues of Eqs. (9.57) and (9.58) (see Appendix I):

$$\hat{\mathbf{L}}\cdot\hat{\mathbf{M}} = \hat{\mathbf{M}}\cdot\hat{\mathbf{L}} = 0; \tag{9.97}$$

$$\hat{\mathbf{M}}^2 = 2m\hat{\mathscr{H}}\left(\hat{\mathbf{L}}^2 + \hbar^2\right) + m^2k^2. \tag{9.98}$$

Notice that the second equation reduces to Eq. (9.58) in the classical limit as Planck's constant tends toward zero: $\hbar \to 0$. This is an example of Bohr's "celebrated" *correspondence principle.*

9.4.2 Conservation laws

The classical conservation laws for the hydrogen atom were described in §9.3.3 and §9.3.4. The *constants of the motion* were found to be the energy E (represented by the Hamiltonian $\mathscr{H}$), the components of the angular momentum vector **L**, and the components of the LRL vector **M**. Using his matrix mechanics approach, Pauli demonstrated that the corresponding quantum entities were constant in time as well. Translating this in operator terminology, we recall from §6.3.2 that the associated operators of conserved quantities in quantum mechanics commute with the Hamiltonian. That is, if an operator $\hat{A}$, corresponding to the observable physical quantity A, commutes with the Hamiltonian (i.e., $\left[\hat{A},\hat{\mathscr{H}}\right] = 0$), then A is a conserved quantity (i.e., $d\langle A\rangle/dt = 0$). It can therefore be postulated that

$$\left[\hat{L}_i, \hat{\mathscr{H}}\right] = 0, \qquad \forall i = 1 \to 3, \tag{9.99}$$

and that

$$\left[\hat{M}_i, \hat{\mathscr{H}}\right] = 0, \qquad \forall i = 1 \rightarrow 3. \tag{9.100}$$

Demonstrating the validity of these assertions is far from trivial, and the actual computations have been relegated to Appendix I for the aficionado.[29] Notice that the previous equations are the quantum analogues of Eqs. (9.38) and (9.50).

Following the discussion in Chapter 6, the operator components of $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$ can be seen as the *generators* of a particular Lie algebra. The corresponding Lie group can then be invoked to rationalize the degeneracies of the hydrogen system. In the following section, we will turn directly to the appropriate Lie group algebra, which incorporates all the operators concerned, and obtain the atomic spectrum of hydrogen as a consequence.[30]

9.5 THE SPECIAL ORTHOGONAL GROUP SO(4)

9.5.1 The generators of SO(4)

The three components of $\hat{\mathbf{M}}$ are generators of infinitesimal transformations in much the same way that the three components of $\hat{\mathbf{L}}$ were regarded as generators of infinitesimal rotations about the three orthogonal axes. We thus proceed by working out the algebra of the six generators $\hat{L}_i$ and $\hat{M}_i$ $(i = 1 \rightarrow 3)$, which consists of all possible commutation relations among the generators. The different elementary commutation relationships on which these results are based are summarized in Appendix I.

Three of these have already been given before and constitute the angular momentum algebra:

$$\left[\hat{L}_i, \hat{L}_j\right] = i\hbar\varepsilon_{ijk}\hat{L}_k. \tag{9.101}$$

Nine additional commutation relations between the components of $\hat{\mathbf{M}}$ and $\hat{\mathbf{L}}$ are given by

$$\left[\hat{M}_i, \hat{L}_j\right] = i\hbar\varepsilon_{ijk}\hat{M}_k, \tag{9.102}$$

establishing $\hat{\mathbf{M}}$ as a *vector operator*. The last three commutators are even more cumbersome to calculate, but lead to the simple result

$$\left[\hat{M}_i, \hat{M}_j\right] = i\hbar\left(-2m\hat{\mathscr{H}}\right)\varepsilon_{ijk}\hat{L}_k. \tag{9.103}$$

The components of $\hat{\mathbf{L}}$ by themselves constitute the closed angular momentum algebra and generate the group SO(3), as we saw in Chapter 5. The $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$ together, however, do not form a closed algebra. Although the commutators in Eq. (9.102) involve only $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$, Eq. (9.103) brings in $\hat{\mathscr{H}}$ as well. However, as McIntosh recounted: “Since the energy is a constant of the motion, too serious a problem does not arise when it appears in the commutation rules, since it can always be replaced by its value, classically, and its eigenvalues, quantum mechanically.”[31] That is, since $\hat{\mathscr{H}}$ is independent of time and commutes with $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$, we can work in the subspace

---

[29] Of course, the energy is conserved as well since $\left[\hat{\mathscr{H}}, \hat{\mathscr{H}}\right] = 0$ obvisously holds.

[30] Pauli went even further and explained the term splitting in an electric field, as observed in the Stark effect. However interesting, we will not consider this aspect any further.

[31] McIntosh, “Symmetry and Degeneracy”, p. 84.

$\mathcal{H}(E)$ of the Hilbert space $\mathcal{H}$ that corresponds to a particular energy eigenvalue E of the Hamiltonian $\hat{\mathscr{H}}$. Then, $\hat{\mathscr{H}}$ may be replaced in Eq. (9.103) by the *restricted* Hamiltonian $\hat{\mathscr{H}}|_{\mathcal{H}(E)} = E$, which for bound states is a negative quantity ($E < 0$).[32] In this way, the energy can be absorbed into a normalization factor. We thus define a new *normalized* Laplace-Runge-Lenz vector $\hat{\mathbf{A}}$ by

$$\hat{\mathbf{A}} \equiv \frac{\hat{\mathbf{M}}}{\sqrt{-2mE}}. \tag{9.104}$$

The commutation relations in Eq. (9.102) hold as well for $\hat{\mathbf{A}}$

$$\begin{aligned}\left[\hat{A}_i, \hat{L}_j\right] &= \left[\frac{\hat{M}_i}{\sqrt{-2mE}}, \hat{L}_j\right] \\ &= \frac{1}{\sqrt{-2mE}}\left[\hat{M}_i, \hat{L}_j\right] \\ &= i\hbar\varepsilon_{ijk}\frac{\hat{M}_k}{\sqrt{-2mE}} \\ \left[\hat{A}_i, \hat{L}_j\right] &= i\hbar\varepsilon_{ijk}\hat{A}_k.\end{aligned} \tag{9.105}$$

The commutators in Eq. (9.103), on the other hand, are replaced by

$$\begin{aligned}\left[\hat{A}_i, \hat{A}_j\right] &= \left[\frac{\hat{M}_i}{\sqrt{-2mE}}, \frac{\hat{M}_k}{\sqrt{-2mE}}\right] \\ &= -\frac{1}{2mE}\left[\hat{M}_i, \hat{M}_j\right] \\ &= -\frac{1}{2mE}i\hbar(-2mE)\,\varepsilon_{ijk}\hat{L}_k \\ \left[\hat{A}_i, \hat{A}_j\right] &= i\hbar\varepsilon_{ijk}\hat{L}_k.\end{aligned} \tag{9.106}$$

The six generators $\hat{L}_i$ and $\hat{A}_i$ clearly constitute a *closed Lie algebra*, as can be seen from Table 9.2. *But which symmetry group are they generating?* One of the clues we can use in answering this question is the number of generators. Among the classical Lie groups are the (special) orthogonal groups in n dimensions $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, and the (special) unitary groups in n dimensions $\mathrm{U}(n)$ and $\mathrm{SU}(n)$. The number of generators for these groups is dependent on the dimensionality n of the group. In Table 9.3 are the number of generators for the Lie groups $\mathrm{SO}(n)$, $\mathrm{U}(n)$, and $\mathrm{SU}(n)$ for $n = 2 \to 5$. Glancing briefly over the numbers in this table reveals that the special orthogonal group in four dimensions, SO(4), is the only Lie group with six generators. Indeed, as we shall see in the next section, the algebra $\{\hat{L}_1, \hat{L}_2, \hat{L}_3, \hat{A}_1, \hat{A}_2, \hat{A}_3\}$ can be identified with the $\mathfrak{so}(4)$

---

[32] However, from a strictly mathematical perspective, this replacement is an ad hoc solution that calls for a deeper algebraic treatment. See, for example, J. Daboul, P. Slodowy, and C. Daboul. "The Hydrogen Algebra as Centerless Twisted Kac-Moody Algebra." *Physics Letters B* 317.3 (1993), pp. 321–328; C. Daboul, J. Daboul, and P. Slodowy. "The Dynamical Algebra of the Hydrogen Atom as a Twisted Loop Algebra." In: *Proceedings of the XX International Colloquium on "Group Theoretical Methods in Physics."* Eds. A. Arima, T. Eguchi, and N. Nakanishi. Singapore: World Scientific, 1995, pp. 175–178.

Table 9.2 Commutation table for the $\hat{\mathbf{L}}$ and $\hat{\mathbf{A}}$ operators in the subspace $\mathcal{H}(E)$ of the Hilbert space $\mathcal{H}$ with $E < 0$. The commutator $[\hat{X}_i, \hat{X}_j]$ $(i, j = 1 \to 6)$ is listed in the ith row and jth column. The commutation table is *skew-symmetric* as a consequence of the fundamental commutation relation $\left[\hat{X}_i, \hat{X}_j\right] = -\left[\hat{X}_j, \hat{X}_i\right]$

	$\hat{L}_1$	$\hat{L}_2$	$\hat{L}_3$	$\hat{A}_1$	$\hat{A}_2$	$\hat{A}_3$
$\hat{L}_1$	0	$i\hbar\hat{L}_3$	$-i\hbar\hat{L}_2$	0	$i\hbar\hat{A}_3$	$-i\hbar\hat{A}_2$
$\hat{L}_2$	$-i\hbar\hat{L}_3$	0	$i\hbar\hat{L}_1$	$-i\hbar\hat{A}_3$	0	$i\hbar\hat{A}_1$
$\hat{L}_3$	$i\hbar\hat{L}_2$	$-i\hbar\hat{L}_1$	0	$i\hbar\hat{A}_2$	$-i\hbar\hat{A}_1$	0
$\hat{A}_1$	0	$i\hbar\hat{A}_3$	$-i\hbar\hat{A}_2$	0	$i\hbar\hat{L}_3$	$-i\hbar\hat{L}_2$
$\hat{A}_2$	$-i\hbar\hat{A}_3$	0	$i\hbar\hat{A}_1$	$-i\hbar\hat{L}_3$	0	$i\hbar\hat{L}_1$
$\hat{A}_3$	$i\hbar\hat{A}_2$	$-i\hbar\hat{A}_1$	0	$i\hbar\hat{L}_2$	$-i\hbar\hat{L}_1$	0

Lie algebra. Having thus uncovered the hidden symmetry of the hydrogen atom as the SO(4) symmetry, we are led to

$$\begin{array}{c} \text{Conservation of} \\ \text{the LRL vector} \end{array} \longleftrightarrow \begin{array}{c} \text{Four-Dimensional} \\ \text{Rotation} \\ \text{Symmetry} \end{array} \longleftrightarrow \begin{array}{c} n^2\text{-fold} \\ \text{Degeneracy} \end{array} \tag{9.107}$$

Before embarking on our voyage to the fourth dimension, we should note one more thing. In the previous treatment, E was required to be negative to ensure a real outcome of $\sqrt{-2mE}$. This requirement is absolutely crucial to have an $\mathfrak{so}(4)$ algebra. It is the operator equivalent of the *boundary condition* in the Schrödinger wave mechanical treatment. According to this boundary condition, the wave function must be *square integrable*, which means that it must vanish asymptotically when the distance between the electron and the proton goes to infinity. This condition keeps the electron attracted to the proton and traps it in the potential well, which gives rise to the bound state. Without this condition, there are no quantized energies, no quantum levels, no hydrogen spectrum.

9.5.2 The $\mathfrak{so}(4)$ Lie algebra

Before we turn to the fourth dimension, let us first consider our familiar three-dimensional Euclidean space $\mathbb{R}^3$, and label the coordinate and momentum vectors as follows:

$$\mathbf{r} = (r_1,\, r_2,\, r_3) \quad \text{and} \quad \mathbf{p} = \big(p_1,\, p_2,\, p_3\big). \tag{9.108}$$

The angular momentum vector in this space consists of three components that from now on may be conveniently labeled by two integer numbers that refer to the coordinate and momentum components involved:

$$L_{ij} \equiv r_i p_j - r_j p_i. \tag{9.109}$$

Notice that, evidently: $L_{ij} = -L_{ji}$. We thus obtain the *natural indices* for the angular momentum vector $\mathbf{L}$:

$$\mathbf{L} = (L_{23},\, L_{31},\, L_{12}). \tag{9.110}$$

Table 9.3 Table with the number of generators for the Lie groups SO(n), U(n), and SU(n) with $n = 2 \rightarrow 5$. Notice that the O(n) groups have the same number of generators as the SO(n) groups.

Dimension n	*Group SO(n)*	*Generators* $n(n-1)/2$	*Group U(n)*	*Generators* n^2	*Group SU(n)*	*Generators* $n^2 - 1$
2	SO(2)	1	U(2)	4	SU(2)	3
3	SO(3)	3	U(3)	9	SU(3)	8
4	SO(4)	6	U(4)	16	SU(4)	15
5	SO(5)	10	U(5)	25	SU(5)	24

Let us now enlarge our space to the four-dimensional Euclidean space $\mathbb{R}^4$ by introducing a fourth component to the coordinate and momentum vectors:

$$\mathbf{r} = (r_1, r_2, r_3, r_4) \quad \text{and} \quad \mathbf{p} = \big(p_1, p_2, p_3, p_4\big). \tag{9.111}$$

In this space, the angular momentum vector $\mathbf{L}$ consists of not less than *six* components,[33] which we assign the natural indices:

$$\mathbf{L} = (L_{23}, L_{31}, L_{12}, L_{14}, L_{24}, L_{34}). \tag{9.112}$$

Turning to the corresponding quantum mechanical operators, the position and momentum operators are seen to obey Heisenberg's canonical commutation relations (see Appendix I):

$$\left[\hat{r}_i, \hat{p}_j\right] = i\hbar\delta_{ij}. \tag{9.113}$$

It is then a straightforward matter to obtain the commutation relations for the six components of $\hat{\mathbf{L}}$ (see Appendix I again):

$$\left[\hat{L}_{ij}, \hat{L}_{ik}\right] = i\hbar\hat{L}_{jk}, \qquad \left[\hat{L}_{ij}, \hat{L}_{kl}\right] = 0 \quad i \neq j \neq k \neq l. \tag{9.114}$$

The results are shown in the commutation Table 9.4. A remarkable observation can now be made. On comparison of the commutation Tables 9.2 and 9.4, we can see that both are *isomorphic* to one another. Worded somewhat differently, we see that the six operators we obtained in the quantum mechanical treatment of the hydrogen atom may, in fact, be identified with the six rotational operators (or *conserved angular momenta*) in four dimensions. The following representation can thus be made:

$$\begin{aligned} \hat{L}_1 &= \hat{L}_{23}; \quad \hat{L}_2 = \hat{L}_{31}; \quad \hat{L}_3 = \hat{L}_{12} \\ \hat{A}_1 &= \hat{L}_{14}; \quad \hat{A}_2 = \hat{L}_{24}; \quad \hat{A}_3 = \hat{L}_{34}. \end{aligned} \tag{9.115}$$

We can easily verify that the commutators in Eq. (9.114) indeed match the commutation relations in Eqs. (9.101), (9.105), and (9.106).

The six generators $\hat{L}_{ij}$ obviously constitute the generalization of the three generators $\hat{L}_1$, $\hat{L}_2$, and $\hat{L}_3$ from three to four dimensions. The group they generate is the

Table 9.4 Commutation table for the generators of the $\mathfrak{so}(4)$ Lie algebra in the basis $\{\hat{L}_{23}, \hat{L}_{31}, \hat{L}_{12}, \hat{L}_{14}, \hat{L}_{24}, \hat{L}_{34}\}$. Note that the $\mathfrak{so}(4)$ Lie algebra is *closed* under commutation and that it is *isomorphic* to the commutation Table 9.2

	$\hat{L}_{23}$	$\hat{L}_{31}$	$\hat{L}_{12}$	$\hat{L}_{14}$	$\hat{L}_{24}$	$\hat{L}_{34}$
$\hat{L}_{23}$	0	$i\hbar\hat{L}_{12}$	$-i\hbar\hat{L}_{31}$	0	$i\hbar\hat{L}_{34}$	$-i\hbar\hat{L}_{24}$
$\hat{L}_{31}$	$-i\hbar\hat{L}_{12}$	0	$i\hbar\hat{L}_{23}$	$-i\hbar\hat{L}_{34}$	0	$i\hbar\hat{L}_{14}$
$\hat{L}_{12}$	$i\hbar\hat{L}_{31}$	$-i\hbar\hat{L}_{23}$	0	$i\hbar\hat{L}_{24}$	$-i\hbar\hat{L}_{14}$	0
$\hat{L}_{14}$	0	$i\hbar\hat{L}_{34}$	$-i\hbar\hat{L}_{24}$	0	$i\hbar\hat{L}_{12}$	$-i\hbar\hat{L}_{31}$
$\hat{L}_{24}$	$-i\hbar\hat{L}_{34}$	0	$i\hbar\hat{L}_{14}$	$-i\hbar\hat{L}_{12}$	0	$i\hbar\hat{L}_{23}$
$\hat{L}_{34}$	$i\hbar\hat{L}_{24}$	$-i\hbar\hat{L}_{14}$	0	$i\hbar\hat{L}_{31}$	$-i\hbar\hat{L}_{23}$	0

---

[33] Given that the indices i and j both range over four values $(1 \to 4)$, and that $i \neq j$ must hold, there are $4 \times 3 = 12$ different combinations. And since $L_{ij} = -L_{ji}$, the number of independent generators halves to six.

proper rotation group or *orthogonal group in four dimensions*, designated SO(4), which is the set of all 4 × 4 real orthogonal matrices with determinant equal to +1 that leave the quadratic form $r_1^2 + r_2^2 + r_3^2 + r_4^2$ invariant. Each of the $\hat{L}_{ij}$ operators generates an infinitesimal rotation in one of the six coordinate planes r_1r_2, r_1r_3, r_2r_3, r_1r_4, r_2r_4, or r_3r_4.[34] This, evidently, does not represent a geometric symmetry of the hydrogen atom, because the fourth components r_4 and p_4 are fictitious and cannot be identified with spatial variables. For this reason, SO(4) is said to describe a *dynamical symmetry* of the hydrogen atom. It does, of course, contain the *geometric symmetry* SO(3) as a subgroup:

$$\mathrm{SO}\,(4) \supset \mathrm{SO}\,(3). \tag{9.116}$$

It is important to note that the $\mathfrak{so}(4)$ generators were obtained by restricting our considerations to *bound states*. For *continuum* (or *scattering*) states, E is positive, and the sign inside the square root of Eq. (9.104) must be changed for $\mathbf{A}$ to be Hermitian. Then, the sign on the *rhs* of Eq. (9.106) is changed (i.e., $\left[\hat{A}_i, \hat{A}_j\right] = -i\hbar\varepsilon_{ijk}\hat{L}_k$), and the identifications in Eq. (9.115) are no longer valid. It turns out that the dynamical symmetry group in this case is isomorphic to the *group of Lorentz transformations* in one time and three space dimensions, rather than to the group of rotations in four space dimensions. This group is denoted SO(3,1) and is especially important in Einstein's theory of special relativity.[35] If, instead, the energy is zero, then the appropriate symmetry group is the Euclidean group in four dimensions, denoted E(4). The relations between the energy E and the corresponding symmetry group are listed in Table 9.5.

The factorization of $\mathfrak{so}(4)$

The structure of the $\mathfrak{so}(4)$ algebra is of an amazing simplicity. By forming the sum and the difference of the $\hat{\mathbf{L}}$ and $\hat{\mathbf{A}}$ operators,

$$\hat{\mathbf{J}}_1 = \frac{1}{2}\left(\hat{\mathbf{L}} + \hat{\mathbf{A}}\right), \quad \hat{\mathbf{J}}_2 = \frac{1}{2}\left(\hat{\mathbf{L}} - \hat{\mathbf{A}}\right), \tag{9.117}$$

Table 9.5 Relation between the total energy E of the system and the corresponding symmetry group.

Energy	*Symmetry group*	*Symbol*
$E < 0$	Orthogonal group	SO(4)
$E = 0$	Euclidean group	E(4)
$E > 0$	Lorentz group	SO(3,1)

---

[34] A rotation in one of the two-dimensional planes r_ir_j can then be denoted by the rotation operator $\hat{U}\left(\theta_{ij}\right) = \exp\left(-\frac{i}{\hbar}\theta_{ij}\hat{L}_{ij}\right)$, with $0 \leq \theta_{ij} \leq 2\pi$.

[35] As John Baez amusingly observed, "Who'd have thought [Einstein's theory of special relativity] was lurking in Newtonian gravity?" Quoted from J. Baez. *Mysteries of the Gravitational 2-Body Problem*. March 16, 2015. Posted on: math.ucr.edu/home/baez/gravitational.html. The two-dimensional analogue of the Lorentz group, denoted SO(2,1) will be discussed in Chapter 11. The SO(3,1) group will also appear in our examination of the periodic system in Chapter 14.

we can see that these two linear combinations of $\hat{\mathbf{L}}$ and $\hat{\mathbf{A}}$ are fully commutative:

$$\left[\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2\right] = 0, \tag{9.118}$$

where the vector notation in the commutator means that any component of the $\hat{\mathbf{J}}$s can be chosen

$$\left[\hat{J}_{1i}, \hat{J}_{2j}\right] = 0, \qquad \forall i, j = 1 \to 3. \tag{9.119}$$

This can be shown as follows:

$$\begin{aligned} \left[\hat{J}_{1i}, \hat{J}_{2j}\right] &= \left[\frac{1}{2}\left(\hat{L}_i + \hat{A}_i\right), \frac{1}{2}\left(\hat{L}_j - \hat{A}_j\right)\right] \\ &= \frac{1}{4}\left(\left[\hat{L}_i, \hat{L}_j\right] + \left[\hat{A}_i, \hat{L}_j\right] - \left[\hat{L}_i, \hat{A}_j\right] - \left[\hat{A}_i, \hat{A}_j\right]\right) \\ &= \frac{1}{4}\left(i\hbar\varepsilon_{ijk}\hat{L}_k + i\hbar\varepsilon_{ijk}\hat{A}_k - i\hbar\varepsilon_{ijk}\hat{A}_k - i\hbar\varepsilon_{ijk}\hat{L}_k\right) \\ \left[\hat{J}_{1i}, \hat{J}_{2j}\right] &= 0. \end{aligned} \tag{9.120}$$

Furthermore, the $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ operators themselves obey the standard commutation relationships for angular momentum:

$$\left[\hat{J}_{1i}, \hat{J}_{1j}\right] = i\hbar\varepsilon_{ijk}\hat{J}_{1k}; \tag{9.121}$$

$$\left[\hat{J}_{2i}, \hat{J}_{2j}\right] = i\hbar\varepsilon_{ijk}\hat{J}_{2k}. \tag{9.122}$$

The proof for Eq. (9.121) goes as follows:

$$\begin{aligned} \left[\hat{J}_{1i}, \hat{J}_{1j}\right] &= \left[\frac{1}{2}\left(\hat{L}_i + \hat{A}_i\right), \frac{1}{2}\left(\hat{L}_j + \hat{A}_j\right)\right] \\ &= \frac{1}{4}\left(\left[\hat{L}_i, \hat{L}_j\right] + \left[\hat{L}_i, \hat{A}_j\right] + \left[\hat{A}_i, \hat{L}_j\right] + \left[\hat{A}_i, \hat{A}_j\right]\right) \\ &= \frac{1}{4}\left(i\hbar\varepsilon_{ijk}\hat{L}_k + i\hbar\varepsilon_{ijk}\hat{A}_k + i\hbar\varepsilon_{ijk}\hat{A}_k + i\hbar\varepsilon_{ijk}\hat{L}_k\right) \\ &= i\hbar\varepsilon_{ijk}\frac{1}{2}\left(\hat{L}_k + \hat{A}_k\right) \\ \left[\hat{J}_{1i}, \hat{J}_{1j}\right] &= i\hbar\varepsilon_{ijk}\hat{J}_{1k}. \end{aligned} \tag{9.123}$$

Eq. (9.122) can be proved along the same lines. The commutation relations in Eqs. (9.121) and (9.122) show that the operators $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ each constitute an $\mathfrak{su}(2)$ algebra, which we denote by $\mathfrak{su}(2)_1$ and $\mathfrak{su}(2)_2$, respectively for further convenience. In view of Eqs. (9.118) and (9.119), these two algebras are completely *decoupled.* This is especially clear from commutation Table 9.6, where the two $\mathfrak{su}(2)$ algebras are separated from each other by two squares of zeroes. The $\mathfrak{so}(4)$ algebra is then seen to be locally isomorphic to the *direct sum* of these two separate algebras:

$$\mathfrak{so}(4) = \mathfrak{su}(2)_1 \oplus \mathfrak{su}(2)_2. \tag{9.124}$$

This is reminiscent of the factorization of the U(n) group as the direct product SU(n) $\otimes$ U(1). The U(3) group, for instance, can be factorized into the direct product SU(3) $\otimes$ U(1), and the same applies to their corresponding Lie algebras: $\mathfrak{u}(3) = \mathfrak{su}(3) \oplus \mathfrak{u}(1)$.

Table 9.6 Commutation table for the generators of the $\mathfrak{so}(4)$ Lie algebra in the basis $\{\hat{J}_{1x},\hat{J}_{1y},\hat{J}_{1z},\hat{J}_{2x},\hat{J}_{2y},\hat{J}_{2z}\}$. The generators $\hat{\mathbf{J}}_1=\left(\hat{J}_{1x},\hat{J}_{1y},\hat{J}_{1z}\right)$ and $\hat{\mathbf{J}}_2=\left(\hat{J}_{2x},\hat{J}_{2y},\hat{J}_{2z}\right)$ each constitute an $\mathfrak{su}(2)$ subalgebra, denoted $\mathfrak{su}(2)_1$ and $\mathfrak{su}(2)_2$, respectively: $\mathfrak{so}(4)=\mathfrak{su}(2)_1\oplus\mathfrak{su}(2)_2$. Both subalgebras are located in a square region, separated from one another by two squares of zeros.

	$\hat{J}_{1x}$	$\hat{J}_{1y}$	$\hat{J}_{1z}$	$\hat{J}_{2x}$	$\hat{J}_{2y}$	$\hat{J}_{2z}$
$\hat{J}_{1x}$	0	$i\hbar\hat{J}_{1z}$	$-i\hbar\hat{J}_{1y}$	0	0	0
$\hat{J}_{1y}$	$-i\hbar\hat{J}_{1z}$	0	$i\hbar\hat{J}_{1x}$	0	0	0
$\hat{J}_{1z}$	$i\hbar\hat{J}_{1y}$	$-i\hbar\hat{J}_{1x}$	0	0	0	0
$\hat{J}_{2x}$	0	0	0	0	$i\hbar\hat{J}_{2z}$	$-i\hbar\hat{J}_{2y}$
$\hat{J}_{2y}$	0	0	0	$-i\hbar\hat{J}_{2z}$	0	$i\hbar\hat{J}_{2x}$
$\hat{J}_{2z}$	0	0	0	$i\hbar\hat{J}_{2y}$	$-i\hbar\hat{J}_{2x}$	0

9.5.3 The Cartan subalgebra of $\mathfrak{so}(4)$

In this section we continue by deriving the *Cartan subalgebra* $\mathfrak{H}$ and corresponding *Weyl diagram* of the $\mathfrak{so}(4)$ Lie algebra. This necessitates a change of basis: from the $\{\hat{J}_{1x},\hat{J}_{1y},\hat{J}_{1z},\hat{J}_{2x},\hat{J}_{2y},\hat{J}_{2z}\}$ basis to the more useful *Cartan-Weyl basis.* To this end, we proceed once again along the threefold path as outlined in Chapter 6, §6.4. Our work will be greatly simplified, however, by the fact that the $\mathfrak{so}(4)$ algebra is composed of two commuting $\mathfrak{su}(2)$ subalgebras, both of which are isomorphic to the familiar angular momentum algebra as described in Chapter 5.

Step 1: The Cartan subalgebra and Cartan generators

In a first step, we identify the maximal subset of mutually commuting generators of the $\mathfrak{so}(4)$ Lie algebra. Since each of the constitutive $\mathfrak{su}(2)$ subalgebras is an algebra of rank 1, we expect to find a maximum of two commuting generators. Indeed, on closer inspection of Table 9.6, we can see that no more than two $\hat{J}$'s can be diagonalized simultaneously—one from the $\mathfrak{su}(2)_1$ algebra and a second one from the $\mathfrak{su}(2)_2$ algebra.

From the nine possible pairs of commuting generators $\{\hat{J}_{1i},\hat{J}_{2j}\}$ $(i,j=x,y,z)$, let us choose the pair $\{\hat{J}_{1z},\hat{J}_{2z}\}$ as our commuting set with

$$\left[\hat{J}_{1z},\hat{J}_{2z}\right]=0. \tag{9.125}$$

The set $\{\hat{J}_{1z},\hat{J}_{2z}\}$ thus forms a basis for the *maximal Abelian Cartan subalgebra* $\mathfrak{H}\subset\mathfrak{so}(4)$, as indicated by the bold zeros in Table 9.7. The operators $\hat{J}_{1z}$ and $\hat{J}_{2z}$ are *Cartan generators*, and the dimension 2 of $\mathfrak{H}$ defines the *rank* of the $\mathfrak{so}(4)$ Lie algebra.

Step 2: Weyl generators

In the second step, we arrange the remaining generators $\hat{J}_{1x}$, $\hat{J}_{1y}$, $\hat{J}_{2x}$, and $\hat{J}_{2y}$ of $\mathfrak{so}(4)$ into linear combinations to form a linearly independent set of *ladder operators* or *Weyl*

Table 9.7 Commutation table for the generators of the $\mathfrak{so}(4)$ Lie algebra in the *Cartan-Weyl basis* $\{\hat{J}_{1z}, \hat{J}_{1+}, \hat{J}_{1-}, \hat{J}_{2z}, \hat{J}_{2+}, \hat{J}_{2-}\}$. The *Cartan subalgebra* $\mathfrak{H} = \{\hat{J}_{1z}, \hat{J}_{2z}\}$ forms a maximal Abelian subalgebra of $\mathfrak{so}(4)$, as indicated by the **zeros**.

	$\hat{J}_{1z}$	$\hat{J}_{1+}$	$\hat{J}_{1-}$	$\hat{J}_{2z}$	$\hat{J}_{2+}$	$\hat{J}_{2-}$
$\hat{J}_{1z}$	$\mathbf{0}$	$\hbar J_{1+}$	$-\hbar\hat{J}_{1-}$	$\mathbf{0}$	0	0
$\hat{J}_{1+}$	$-\hbar J_{1+}$	0	$2\hbar\hat{J}_{1z}$	0	0	0
$\hat{J}_{1-}$	$\hbar\hat{J}_{1-}$	$-2\hbar\hat{J}_{1z}$	0	0	0	0
$\hat{J}_{2z}$	$\mathbf{0}$	0	0	$\mathbf{0}$	$\hbar\hat{J}_{2+}$	$-\hbar\hat{J}_{2-}$
$\hat{J}_{2+}$	0	0	0	$-\hbar\hat{J}_{2+}$	0	$2\hbar\hat{J}_{2z}$
$\hat{J}_{2-}$	0	0	0	$\hbar\hat{J}_{2-}$	$-2\hbar\hat{J}_{2z}$	0

generators. Since $\mathfrak{so}(4)$ is composed of two $\mathfrak{su}(2)$ subalgebras, this problem reduces to finding the *raising* and *lowering* operators for each $\mathfrak{su}(2)$ algebra. Analogous to the angular momentum algebra, we thus define

$$\begin{aligned} \hat{J}_{1+} &\equiv \hat{J}_{1x} + i\hat{J}_{1y}; \qquad & \hat{J}_{1-} &\equiv \hat{J}_{1x} - i\hat{J}_{1y}; \\ \hat{J}_{2+} &\equiv \hat{J}_{2x} + i\hat{J}_{2y}; \qquad & \hat{J}_{2-} &\equiv \hat{J}_{2x} - i\hat{J}_{2y}. \end{aligned} \tag{9.126}$$

The four Weyl generators in Eq. (9.126), along with the two Cartan generators $\hat{J}_{1z}$ and $\hat{J}_{2z}$, form the *Cartan-Weyl basis* for the $\mathfrak{so}(4)$ algebra:

$$\{\hat{J}_{1z}, \hat{J}_{2z}, \hat{J}_{1+}, \hat{J}_{1-}, \hat{J}_{2+}, \hat{J}_{2-}\}. \tag{9.127}$$

By virtue of their function as step operators, the Weyl elements act as *eigenoperators* of the Cartan generators. That is, they satisfy the general commutation relation

$$\left[\hat{H}_i, \hat{E}_\alpha\right] = \alpha_i \hat{E}_\alpha, \qquad \forall i = 1, 2,\ \alpha = 1 \to 4, \tag{9.128}$$

where we have denoted the Cartan generators by the general symbol $\hat{H}_i$ $(i = 1, 2)$ and the Weyl generators by the symbol $\hat{E}_\alpha$ $(\alpha = 1 \to 4)$. The different eigenvalues α_i are referred to as the *roots* of $\hat{E}_\alpha$ with respect to $\hat{H}_i$; their values can be read in Table 9.7.

Step 3: Casimir invariants

Finally, following Racah's theorem and the fact that the $\mathfrak{so}(4)$ Lie algebra is of rank 2, we expect to find two independent *Casimir invariants* $\hat{C}_\mu$ that commute with all the generators of the $\mathfrak{so}(4)$ algebra, including—in particular—the two Cartan elements $\hat{H}_i$:

$$\left[\hat{C}_\mu, \hat{H}_i\right] = 0, \qquad \forall \mu = 1 \to 2,\ i = 1 \to 2. \tag{9.129}$$

Once again, the problem of finding the $\hat{C}_\mu$s reduces to finding the Casimir invariant of each $\mathfrak{su}(2)$ subalgebra separately. Proceeding in analogy with the angular momentum algebra, we define the $\hat{C}_\mu$s as follows:

$$\hat{C}_1 \equiv \hat{\mathbf{J}}_1^2 = \frac{1}{4}\left(\hat{\mathbf{L}} + \hat{\mathbf{A}}\right)^2; \tag{9.130}$$

$$\hat{C}_2 \equiv \hat{\mathbf{J}}_2^2 = \frac{1}{4}\left(\hat{\mathbf{L}} - \hat{\mathbf{A}}\right)^2. \tag{9.131}$$

In view of Eqs. (9.125) and (9.129), there exists a complete set of states in the subspace $\mathcal{H}(E)$ of $\mathcal{H}$ that includes simultaneous eigenstates of $\hat{\mathbf{J}}_1^2$, $\hat{\mathbf{J}}_2^2$, $\hat{J}_{1z}$, and $\hat{J}_{2z}$. If we denote their eigenvalues by j_1, j_2, m_{j1}, and m_{j2}, respectively, we can represent the common eigenstates by the ket

$$\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle. \tag{9.132}$$

Eigenvalues of the Casimir operators $\hat{\mathbf{J}}_1^2$ and $\hat{\mathbf{J}}_2^2$ are then characterized by integer or half-integer angular quantum numbers $j_1, j_2 \in \left\{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\right\}$, such that $\mathbf{J}_1^2 = j_1\left(j_1+1\right)\hbar^2$ and $\mathbf{J}_2^2 = j_2\left(j_2+1\right)\hbar^2$. That is,

$$\hat{\mathbf{J}}_1^2\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle = j_1\left(j_1+1\right)\hbar^2\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle; \tag{9.133}$$

$$\hat{\mathbf{J}}_2^2\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle = j_2\left(j_2+1\right)\hbar^2\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle. \tag{9.134}$$

Each SO(4) manifold, denoted $\left(j_1, j_2\right)$, contains $\left(2j_1+1\right)\left(2j_2+1\right)$ components according to the standard relationships:

$$\hat{J}_{1z}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle = \hbar m_{j1}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle, \tag{9.135}$$

$$\hat{J}_{2z}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle = \hbar m_{j2}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle, \tag{9.136}$$

with $m_{j1} \in \left\{-j_1, -j_1+1, \ldots, j_1-1, j_1\right\}$ and $m_{j2} \in \left\{-j_2, -j_2+1, \ldots, j_2-1, j_2\right\}$. The m_{j1} and m_{j2} eigenvalues are referred to as the *weights* of the Cartan generators $\hat{J}_{1z}$ and $\hat{J}_{2z}$.

SO(4) Weyl diagrams

The *Weyl diagrams* of SO(4) are based on the Cartan subalgebra $\mathfrak{H}$ of the two operators $\hat{J}_{1z}$ and $\hat{J}_{2z}$, which form the basis for a two-dimensional orthogonal frame. In these diagrams, the weights m_{j1} and m_{j2} (in units of $\hbar$) are used as coordinates to plot (and thus label/distinguish) every state of the SO(4) multiplet in the J_{1z}–J_{2z}-plane; that is, they form the components of a two-dimensional *weight vector* $\mathbf{h} = \left(m_{j1}, m_{j2}\right)$, which points from the origin to the state $\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle$. Every multiplet is characterized by the constant eigenvalues j_1 and j_2 of the Casimir invariants. An SO(4) manifold $\left(j_1, j_2\right)$ is thus represented in the Weyl diagram by $\left(2j_1+1\right)\left(2j_2+1\right)$ points. An example is given in Figure 9.8 for the $(1,1)$ manifold.

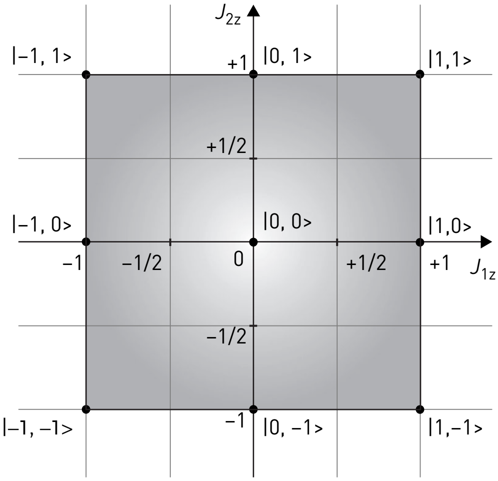


FIGURE 9.8 Weyl diagram of the $(1,1)$ manifold of the SO(4) group. Each point in the diagram corresponds to a given state of the SO(4) multiplet, as indicated by the kets $\left|m_{j1}, m_{j2}\right\rangle$.

When a Weyl diagram has been set up as in Figure 9.8, the Weyl generators enable us to move between the states of the SO(4) manifold by shifting the eigenvalues m_{j1} and m_{j2} of any ket $\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle$ by an amount given by the roots α_1 and α_2 of that Weyl generator with respect to the Cartan generators $\hat{J}_{1z}$ and $\hat{J}_{2z}$:

$$\hat{E}_\alpha \left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \rightarrow \left|j_1, j_2, m_{j1}+\alpha_1, m_{j2}+\alpha_2\right\rangle. \tag{9.137}$$

Let us examine this by considering the action of the shift operator $\hat{J}_{1+}$ on the state $\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle$. By Eqs. (9.128), (9.135), and the results in Table 9.7, we have

$$\begin{aligned}
\hat{J}_{1z}\hat{J}_{1+}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle &= \left(\left[\hat{J}_{1z}, \hat{J}_{1+}\right] + \hat{J}_{1+}\hat{J}_{1z}\right)\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \\
&= \left(\hbar\hat{J}_{1+} + \hbar m_{j1}\hat{J}_{1+}\right)\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \\
&= \hbar\left(m_{j1}+1\right)\hat{J}_{1+}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle.
\end{aligned} \tag{9.138}$$

The operator $\hat{J}_{1+}$ is seen to raise the eigenvalue $\hbar m_{j1}$ by an amount $+\hbar$, which is the root of $\hat{J}_{1+}$ with respect to $\hat{J}_{1z}$ according to Table 9.7. Similarly, using Eq. (9.136), we obtain

$$\begin{aligned}
\hat{J}_{2z}\hat{J}_{1+}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle &= \left(\left[\hat{J}_{2z}, \hat{J}_{1+}\right] + \hat{J}_{1+}\hat{J}_{2z}\right)\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \\
&= \left(0 + \hbar m_{j2}\hat{J}_{1+}\right)\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \\
&= \hbar m_{j2}\hat{J}_{1+}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle,
\end{aligned} \tag{9.139}$$

where the step operator $\hat{J}_{1+}$ leaves the eigenvalue $\hbar m_{j2}$ untouched. Summarizing the previous results yields

$$\hat{J}_{1+}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \rightarrow \left|j_1, j_2, m_{j1}+1, m_{j2}\right\rangle, \tag{9.140}$$

in units of $\hbar$. The actions of the other three Weyl generators can be deduced analogously and are given by

$$\hat{J}_{1-}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \rightarrow \left|j_1, j_2, m_{j1}-1, m_{j2}\right\rangle; \tag{9.141}$$

$$\hat{J}_{2+}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \rightarrow \left|j_1, j_2, m_{j1}, m_{j2}+1\right\rangle; \tag{9.142}$$

$$\hat{J}_{2-}\left|j_1, j_2, m_{j1}, m_{j2}\right\rangle \rightarrow \left|j_1, j_2, m_{j1}, m_{j2}-1\right\rangle. \tag{9.143}$$

SO(4) root diagram

Let us illustrate these actions graphically in a *root diagram*. To this end, we take the roots α_1 and α_2 of every Weyl element $\hat{E}_\alpha$ as the components of a two-dimensional *root vector* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ and position them in the two-dimensional *weight space* formed by the J_{1z}–J_{2z}-plane. This yields the root diagram of the $\mathfrak{so}(4)$ Lie algebra, as depicted in Figure 9.9, where the different root vectors $\boldsymbol{\alpha}$ are denoted by their corresponding Weyl operator symbol $\hat{E}_\alpha$ for simplicity's sake.

Clearly, the $\hat{J}_{1-}$ and $\hat{J}_{1+}$ operators enable us to move one step to the *left* and to the *right*, respectively, whereas one can move *up* and *down* by the $\hat{J}_{2+}$ and $\hat{J}_{2-}$ operators, respectively. The states of an SO(4) multiplet can thus be interconverted by the repeated action of these ladder operators, an example of which is given in the next section.

We can also position the Cartan generators in the SO(4) root diagram. However, in view of their commuting property (Eq. (9.125)), the roots of the Cartan generators

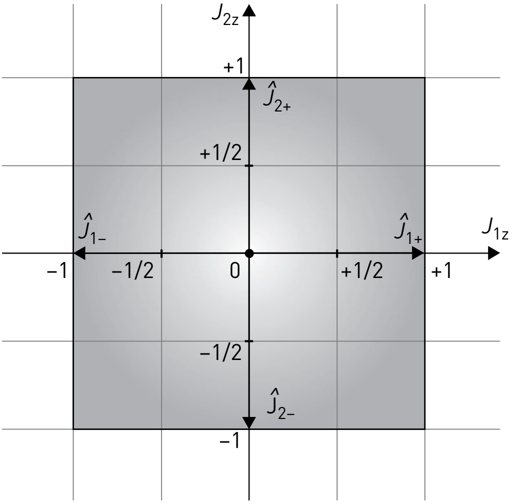


FIGURE 9.9 Root diagram of the $\mathfrak{so}(4)$ Lie algebra. The action of every Weyl generator is shown in the J_{1z}–J_{2z}-plane. The Cartan generators $\hat{J}_{1z}$ and $\hat{J}_{2z}$ are positioned at the origin of the Weyl diagram.

$\hat{J}_{1z}$ and $\hat{J}_{2z}$ are all zero; both Cartan elements are therefore located at the origin of the graph.

With the $\mathfrak{so}(4)$ generators positioned in Figure 9.9, it becomes evident that they correspond to two different manifolds. The $\hat{\mathbf{J}}_1$ operators $\hat{J}_{1z}$, $\hat{J}_{1+}$, and $\hat{J}_{1-}$ (forming the $\mathfrak{su}(2)_1$ subalgebra) correspond to a $(1,0)$ manifold, whereas the $\hat{\mathbf{J}}_2$ operators $\hat{J}_{2z}$, $\hat{J}_{2+}$, and $\hat{J}_{2-}$ are seen to form a $(0,1)$ manifold in the root diagram (corresponding to the $\mathfrak{su}(2)_2$ subalgebra).

Shattering the four-dimensional symmetry

We saw that SO(3) is a subgroup of SO(4) in Eq. (9.116). When reducing the rotation group in four dimensions to the spherical symmetry group in real space, the following *symmetry breaking* occurs:

$$\mathrm{SO}(4) \supset \mathrm{SO}(3). \tag{9.144}$$

This corresponds to a transition from the $\mathfrak{so}(4)$ Lie algebra to the $\mathfrak{so}(3)$ subalgebra, which is described by the standard angular momentum operators $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$. Interestingly, this appears as a *projection onto the diagonal direction* in the root diagram,[36] since

$$\hat{L}_z = \hat{J}_{1z} + \hat{J}_{2z}, \tag{9.145}$$

in view of Eq. (9.117). It can be similarly shown that

$$\hat{L}_+ = \hat{J}_{1+} + \hat{J}_{2+} \quad \text{and} \quad \hat{L}_- = \hat{J}_{1-} + \hat{J}_{2-}. \tag{9.146}$$

Figure 9.10 illustrates the breaking of the sextet of SO(4) generators into two triplets of SO(3) generators.

---

[36] For more information on this *projection theorem*, see A. P. Stone. “Semisimple Subgroups of Semisimple Groups.” *Journal of Mathematical Physics* 11.1 (1970), pp. 29–38.

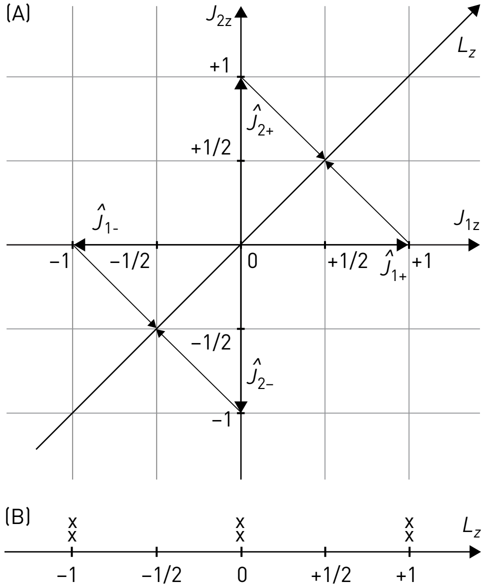


FIGURE 9.10 (A) Root diagram showing the reduction of the $\mathfrak{so}(4)$ Lie algebra to the $\mathfrak{so}(3)$ subalgebra via a *projection* on the diagonal L_z-axis. (B) The breaking of the sextet of SO(4) generators results in two triplets of SO(3) generators, as shown by the x's on the horizontal L_z-axis.

9.6 THE ORIGIN OF ACCIDENTAL DEGENERACIES

9.6.1 Energy levels of the hydrogen atom

The energy eigenvalues for the hydrogen atom can now be found with practically no further effort. Taking the sum and difference of the two Casimir operators $\hat{\mathbf{J}}_1^2$ and $\hat{\mathbf{J}}_2^2$ yields, by Eq. (9.130):

$$\begin{aligned}\hat{\mathbf{J}}_1^2+\hat{\mathbf{J}}_2^2 &= \frac{1}{4}\left(\hat{\mathbf{L}}+\hat{\mathbf{A}}\right)^2+\frac{1}{4}\left(\hat{\mathbf{L}}-\hat{\mathbf{A}}\right)^2\\ &= \frac{1}{4}\left(\hat{\mathbf{L}}^2+\hat{\mathbf{A}}^2+\hat{\mathbf{L}}\cdot\hat{\mathbf{A}}+\hat{\mathbf{A}}\cdot\hat{\mathbf{L}}\right)+\frac{1}{4}\left(\hat{\mathbf{L}}^2+\hat{\mathbf{A}}^2-\hat{\mathbf{L}}\cdot\hat{\mathbf{A}}-\hat{\mathbf{A}}\cdot\hat{\mathbf{L}}\right)\\ &= \frac{1}{2}\left(\hat{\mathbf{L}}^2+\hat{\mathbf{A}}^2\right);\end{aligned} \tag{9.147}$$

$$\begin{aligned}\hat{\mathbf{J}}_1^2-\hat{\mathbf{J}}_2^2 &= \frac{1}{4}\left(\hat{\mathbf{L}}+\hat{\mathbf{A}}\right)^2-\frac{1}{4}\left(\hat{\mathbf{L}}-\hat{\mathbf{A}}\right)^2\\ &= \frac{1}{4}\left(\hat{\mathbf{L}}^2+\hat{\mathbf{A}}^2+\hat{\mathbf{L}}\cdot\hat{\mathbf{A}}+\hat{\mathbf{A}}\cdot\hat{\mathbf{L}}\right)-\frac{1}{4}\left(\hat{\mathbf{L}}^2+\hat{\mathbf{A}}^2-\hat{\mathbf{L}}\cdot\hat{\mathbf{A}}-\hat{\mathbf{A}}\cdot\hat{\mathbf{L}}\right)\\ &= \frac{1}{2}\left(\hat{\mathbf{L}}\cdot\hat{\mathbf{A}}+\hat{\mathbf{A}}\cdot\hat{\mathbf{L}}\right).\end{aligned} \tag{9.148}$$

In view of Eq. (9.97), which is a special property of the LRL vector for the hydrogen atom, the scalar products $\hat{\mathbf{L}}\cdot\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\cdot\hat{\mathbf{L}}$ vanish, yielding for Eq. (9.146):

$$\hat{\mathbf{J}}_1^2-\hat{\mathbf{J}}_2^2=0 \quad \Rightarrow \quad \hat{\mathbf{J}}_1^2=\hat{\mathbf{J}}_2^2. \tag{9.149}$$

This implies that the hydrogen levels are characterized by the *equality* of the two Casimir operators. Hence, we get

$$j_1=j_2. \tag{9.150}$$

The eigenstates of the hydrogen atom thus always correspond to a (j,j) SO(4) manifold. Note that the *other* manifolds of the SO(4) group are of no further interest to us. Given Eq. (9.149), we rewrite Eq. (9.147) as

$$2\hat{\mathbf{J}}_1^2 = \frac{1}{2}\left(\hat{\mathbf{L}}^2 + \hat{\mathbf{A}}^2\right) \quad \text{or} \quad 4\hat{\mathbf{J}}_1^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{A}}^2. \tag{9.151}$$

The eigenvalues of this operator are then

$$4\mathbf{J}_1^2 = 4j\left(j+1\right)\hbar^2 \quad \text{for} \quad j = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots. \tag{9.152}$$

Rewriting Eq. (9.98) as

$$\frac{\hat{\mathbf{M}}^2}{2mE} = \hat{\mathbf{L}}^2 + \hbar^2 + \frac{mk^2}{2E}, \tag{9.153}$$

we obtain, together with Eq. (9.104),

$$\hat{\mathbf{L}}^2 + \hat{\mathbf{A}}^2 = \hat{\mathbf{L}}^2 - \frac{\hat{\mathbf{M}}^2}{2mE} = -\frac{mk^2}{2E} - \hbar^2. \tag{9.154}$$

Combining this relation with Eqs. (9.151) and (9.152) gives

$$4j\left(j+1\right)\hbar^2 = -\frac{mk^2}{2E} - \hbar^2, \tag{9.155}$$

or

$$\left(4j^2 + 4j + 1\right)\hbar^2 = -\frac{mk^2}{2E}, \tag{9.156}$$

which finally yields the spectral result:

$$E_j = -\frac{mk^2}{2\hbar^2\left(2j+1\right)^2} \quad \text{for} \quad j = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots. \tag{9.157}$$

The negative eigenvalues of the hydrogenic Hamiltonian $\hat{\mathscr{H}}$ are in a *one-to-one correspondence* with the eigenvalues $j\left(j+1\right)\hbar^2$ of $\hat{\mathbf{J}}_1^2$. Notice also that Eq. (9.157) agrees with the wave equation result with which we started this chapter:

$$E_n = -\frac{mZ^2e^4}{8h^2{\varepsilon_0}^2 n^2} \quad \text{for} \quad n = 1,\ 2,\ 3,\ 4,\ \ldots, \tag{9.158}$$

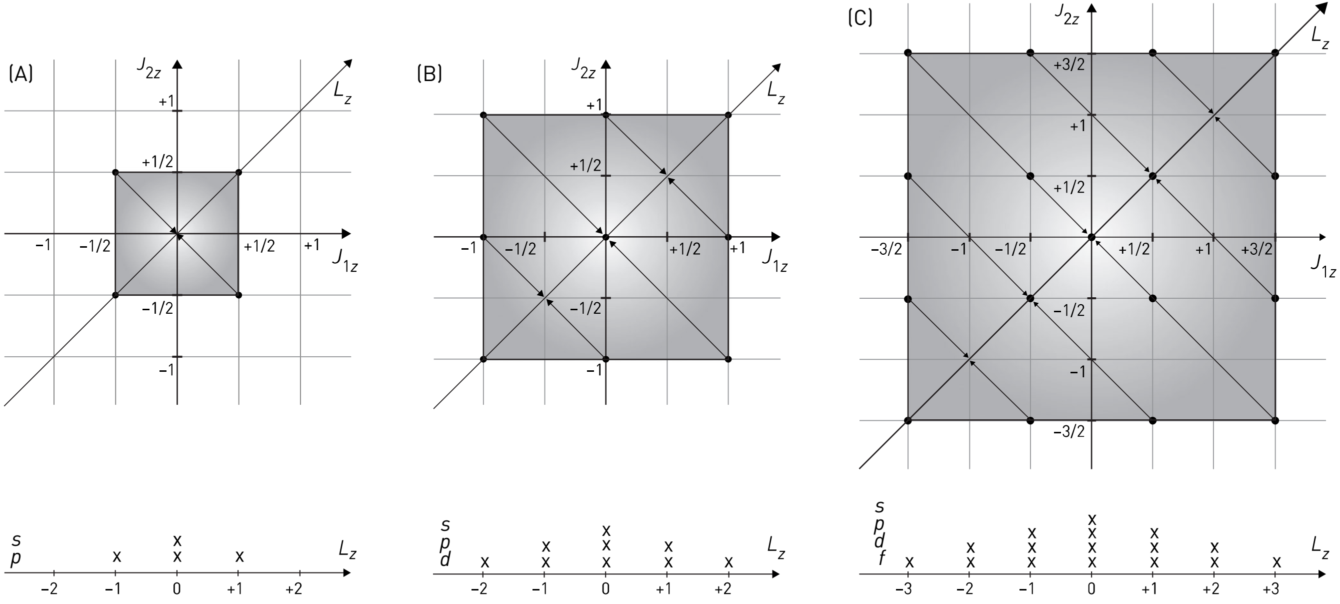
if we remember that $k = Ze^2/4\pi\varepsilon_0$, $\hbar = h/2\pi$, and make the natural identification

$$n \equiv 2j + 1, \tag{9.159}$$

which gives the principal quantum number n the sequence of values 1, 2, 3, ….

It is important to note that there is no objection to using half odd-integer values for j. The only physical restriction is that $\mathbf{L}^2 = l\left(l+1\right)\hbar^2$ has only integer values of l. But, since $\mathbf{L} = \mathbf{J}_1 + \mathbf{J}_2$ from Eq. (9.117), the triangle rule shows that l can have any value ranging from $2j = n-1$ down to $j - j = 0$, by integer steps. Thus, l not only is restricted to integer values, but it also has the correct range of values with respect to the total quantum number n. The degeneracy of this energy level is also given correctly since J_{1z} and J_{2z} can each have $2j+1 = n$ independent eigenvalues, and there are therefore n^2 possible states altogether.

In Figure 9.11, we see the SO(4) Weyl diagrams for the first few eigenlevels of hydrogen, indicating their orbital quantum characteristics by tracking the projections along the diagonal. For $n = 2$, for example, we have $j = (2-1)/2 = 1/2$ by Eq. (9.159). The $(1/2, 1/2)$ manifold thus corresponds to the $n = 2$ level. Its multiplicity equals



(1/2, 1/2) manifold

(1, 1) manifold

(3/2, 3/2) manifold

FIGURE 9.11 SO(4) Weyl diagrams for the first few eigenlevels of hydrogen. (A) For $n = 2$, $j = 1/2$. We therefore associate the $(1/2, 1/2)$ manifold with the $n = 2$ level with multiplicity $n^2 = 4$. A projection of these degenerate states on the diagonal L_z-axis result in a *singlet* ($2s$) and a *triplet* ($2p_x$, $2p_y$, and $2p_z$), as indicated by the x's on the horizontal L_z-axis. (B) Similarly, for $n = 3$, $j = 1$ and a $(1, 1)$ manifold is obtained. The nine degenerate states result in one *singlet* ($3s$), one *triplet* ($3p_x$, $3p_y$, and $3p_z$), and one *quintet* ($3d_{xy}$, $3d_{xz}$, $3d_{yz}$, $3d_{x^2+y^2}$, and $3d_{z^2}$) when projected along the diagonal. (C) For $n = 4$, $j = 3/2$. The $(3/2, 3/2)$ manifold contains $n^2 = 16$ states that break into a *singlet* ($4s$), a *triplet* ($4p_x$, $4p_y$, and $4p_z$), a *quintet* ($4d_{xy}$, $4d_{xz}$, $4d_{yz}$, $4d_{x^2+y^2}$, and $4d_{z^2}$), and a *septet* (f_{z^3}, f_{xz^2}, f_{yz^2}, f_{xyz}, $f_{z(x^2-y^2)}$, $f_{x(x^2-3y^2)}$, and $f_{y(3x^2-y^2)}$).

$n^2 = 4$. A projection of these four points on the diagonal L_z-axis results in a *singlet* and a *triplet*, which we identify as the $2s$-orbital and the three $2p$-orbitals ($2p_x$, $2p_y$, and $2p_z$). Similarly, for $n = 3$, $j = (3-1)/2 = 1$. The $(1,1)$ manifold has a multiplicity of $n^2 = 9$. Projecting the nine degenerate states onto the diagonal results in one *singlet* ($3s$), one *triplet* ($3p_x$, $3p_y$, and $3p_z$), and one *quintet* ($3d_{xy}$, $3d_{xz}$, $3d_{yz}$, $3d_{x^2+y^2}$, and $3d_{z^2}$). When the $(3/2, 3/2)$ multiplet on the *rhs* of Figure 9.11 is broken to the SO(3) symmetry, a *septet* is formed as well, corresponding to the $4f$-orbitals. It should be noted that, within an SO(4) multiplet, all the orbitals are related to one another. A $4s$-orbital can thus be converted into a $4d_{xz}$-orbital, for example, and a $3p_z$-orbital into a $3d_{z^2}$-orbital.

9.6.2 History of the LRL vector (II)

Pauli's algebraic approach

As noted in the introduction to this chapter, the year 1926 was a milestone in the history of quantum mechanics. Not only did Schrödinger's wave mechanics see the light, but also Heisenberg's matrix mechanics gained widespread acceptance among physicists. One of the principal reasons for this was Wolfgang Pauli's successful derivation of the hydrogen spectrum in purely algebraic terms.[37]

When drawing the correspondence between the classical Kepler system and the quantum mechanical Coulomb system, Pauli constructed a set of operators $\hat{\mathscr{H}}$, $\hat{\mathbf{L}}$, and $\hat{\mathbf{M}}$ that mirrored the constants of the motion $\mathscr{H}$, $\mathbf{L}$, and $\mathbf{M}$ of the Kepler problem. As the editors of Pauli's *Collected Works* observed: "For Pauli, the invariants in physics were the symbols of ultimate truth which must be attained by penetrating through the accidental details of things."[38] With the operators at hand, Pauli set out to apply the operational methods as laid down in Heisenberg's matrix mechanics, and thus accounted for the n^2 multiplicity of the experimentally observed hydrogen spectrum.

It seems that Pauli's familiarity with the work of German physicist Wilhelm Lenz (1988–1957) had provided the impetus to exploit the powers of the LRL vector.[39] Lenz had published an article in 1924 in which he applied the LRL vector to a quantum mechanical treatment of the hydrogen atom to derive the classical Kepler trajectories.[40] This work was still performed within the framework of the *old quantum theory*, in which an exquisite mixture of both classical and quantum ideas were cross-fertilizing one another. As Lenz had referred to Runge's famous treatise on *Vektoranalysis* as the source for the "little known" vector, the vector **M** came to be known as the Runge-Lenz vector.

---

[37] See Pauli, Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik," translated as Pauli, "On the Hydrogen Spectrum from the Standpoint of the New Quantum Mechanics."

[38] R. Krong and V. F. Weisskopf. "Preface." In: W. Pauli. *Collected Scientific Papers by Wolfgang Pauli*, vol. 1. Eds. R. Kronig and V. F. Weisskopf. New York: Interscience Publishers, 1964, p. viii.

[39] Pauli had served as Lenz's assistant at Hamburg in 1922, and he discussed Lenz's work at length in his monograph on the old *Quantentheorie*. See W. Pauli. "Quantentheorie." In: *Handbuch der Physik*, vol. 23. Eds. H. Geiger and K. Scheel. Berlin: Springer, 1926, pp. 1–278.

[40] W. Lenz. Über den Bewegungsverlauf und die Quantenzustände der gestörten Keplerbewegung." *Zeitschrift für Physik* 24.1 (1924), pp. 197–207.

Fock's analytic approach

At the time, Pauli did not yet have the necessary group theoretical tools at hand to reflect on the symmetries of the hydrogen atom.[41] But, in 1935, a groundbreaking paper appeared in the *Zeitschrift für Physik*.[42] Written by Soviet physicist Vladimir Aleksandrovich Fock (1898–1974) and titled "*Zur Theorie des Wasserstoffatoms*," the article started with these authoritative words:

> It has long been known that the energy levels of the hydrogen atom are degenerate with respect to their azimuthal quantum number l; one speaks occasionally of an "accidental" degeneracy. But every degeneracy in the eigenvalues is in correspondence with a transformation group: as for example, the degeneracy with respect to the magnetic quantum number m is in connection with the ordinary rotation group. However, the group which corresponds to the "accidental" degeneracy of the hydrogen atom was, up to now, unknown.[43]

Fock's rationalization of the accidental degeneracies in the hydrogen spectrum revolved around the central idea of a *supersymmetry* —hidden in the fourth dimension. In a nutshell, Fock represented the hydrogenic wave equations in *momentum space* (rather than position space).[44] These wave functions were then *projected stereographically* from our familiar three dimensions onto the surface of a four-dimensional hypersphere (or 3-sphere $\mathbb{S}^3$, see also §9.6.4). This is where all the magic happened: the projected wavefunctions were seen to satisfy Schrödinger's equation for a free particle on $\mathbb{S}^3$; that is, the description of the bound states of the electron in the presence of a Coulomb field turned out to be mathematically equivalent to the motion of a free point particle, confined to the surface of a 3-sphere. Both systems were, in other words, *isomorphic* to one another.[45]

The connection between Fock's analytic approach and Pauli's algebraic approach was drawn six months later in 1936 by Valentine Bargmann (1908–1989), who noticed that the operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$ in Pauli's article functioned as the *generators of infinitesimal rotations* of the four-dimensional hypersphere, giving rise to the SO(4) group as outlined by Fock.[46] Since the angular momentum components were known to generate the familiar three-dimensional rotation group SO(3), the components of

---

[41] He dealt with the four-dimensional rotation group thirty years later in a series of lectures on continuous groups and reflections in quantum mechanics, delivered in Copenhagen, Zürich and CERN.

[42] V. Fock. "Zur Theorie des Wasserstoffatoms." In: *Zeitschrift für Physik* 98.3 (1935), pp. 145–154.

[43] Ibid., p. 145.

[44] In general, a wave function $\Phi(p)$ in momentum space is related to the same function $\Psi(x)$ in position space via a *Fourier transform*.

[45] The scattering states of hydrogen, with $E > 0$, are mapped onto a *hyperboloid*, which is preserved under the symmetry transformations of the homogeneous Lorentz group SO(3,1). For $E = 0$, the wave functions are projected onto a *hyperplane*, as described by the Euclidean group in four dimensions E(4).

[46] See V. Bargmann. "Zur Theorie des Wasserstoffatoms." *Zeitschrift für Physik* 99.7 (1936), pp. 576–582. In a postscript to his paper, Bargmann noted that Oskar Klein (1894–1977) and Lamek Hulthén (1909–1995) had made similar comments in 1933 by indicating that the components of $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$ formed a Lie algebra isomorphic to $\mathfrak{so}(4)$. See also L. Hulthén. "Über die quantenmechanische Herleitung der Balmerterme." In: *Zeitschrift für Physik* 86 (1933), pp. 21–23.

the LRL vector were deemed responsible for the additional symmetry and accidental degeneracy of the hydrogen atom.

Fock's analytic treatment of the hydrogen atom—although mysterious and incomprehensible at first sight—can be motivated by a consideration of the analogous Kepler problem in classical physics. This necessitates two crucial steps. First, a shift has to be made from position space to momentum space. This will be done in §9.6.3. By considering the Kepler problem in momentum space, the importance of the LRL vector (as noted by Bargmann) will also become more obvious. In a second step, the Kepler problem will be analyzed in terms of hyperspheres and stereographic projections (see §9.6.4). Finally, an attempt will be made to explain Fock's research in a more intuitive way (see §9.6.5).

9.6.3 Hodographs in momentum space

We have seen in §9.3.6 how Kepler was forced to overthrow the Platonic ideal of circular orbits (as expounded in his "Mysterium Cosmographicum")[47] to introduce the elliptical trajectories of the planets. And yet, Kepler had been right, in a certain sense, to maintain that the planets move in circles around the sun; one just has to consider the Kepler problem in *momentum space* (or *velocity space*, if you like), rather than in *position space*.

Let us see how this works. For each position vector **r** in configuration space, there is a corresponding momentum vector **p** directed *tangentially* to the spatial trajectory. As the planet orbits the sun, both the direction and the magnitude of **p** are seen to change continuously (Figure 9.12A). By plotting all the momentum vectors from a

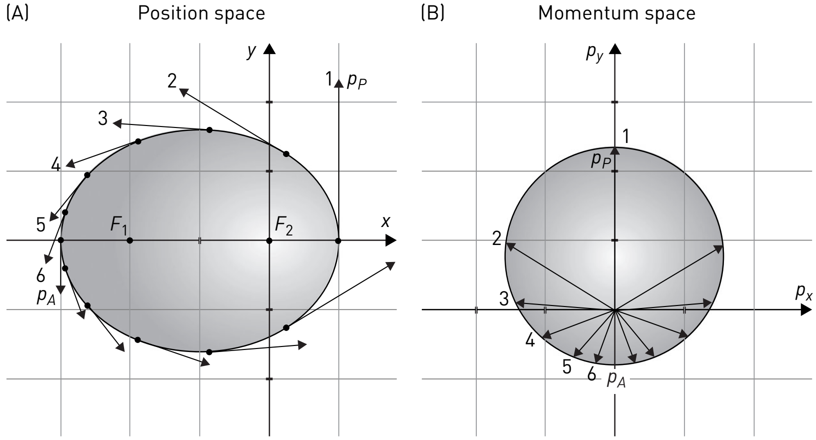


FIGURE 9.12 (A) The elliptical orbit of a planet in *position space*. At each position **r** on the orbit, there is a momentum vector **p** directed *tangentially* to the spatial trajectory, as illustrated for a number of momentum vectors **p** at equal time intervals. According to Kepler's second law, the magnitude of the momentum vector is greatest at the perihelion *P* (see $\mathbf{p}_P$). The vector **p** reaches its smallest value at the aphelion *A* (see $\mathbf{p}_A$). (B) When the momentum vectors are plotted from a common origin *O*, the tips of the vectors generate a *circle* in *momentum space*. This is called a *hodograph*. Notice that each vector in momentum space is *parallel* to the corresponding vector in position space and is of equal magnitude. (Illustration adapted from E. I. Butikov. "The Velocity Hodograph for an Arbitrary Keplerian Motion." *European Journal of Physics* 21 (2000), p 2).

---

[47] J. Kepler. *Mysterium Cosmographicum*. 1596.

common origin, the tips of these vectors trace out a curve in momentum space, which (quite miraculously) turns out to be a perfect *circle* under an inverse square central force, as depicted in Figure 9.12B.

This remarkable fact was first discovered by Sir William Rowan Hamilton (1805–1865), who called the orbit in momentum space a *hodograph*, from the Greek words *οδός* for *way* and *γράφειν* for *to write* or *to describe*. His "Law of the Circular Hodograph" was first communicated before the Royal Irish Academy on December 14, 1846, and was subsequently published in the *Proceedings of the Royal Irish Academy* in 1847.[48] In his demonstration of the circularity of the hodographic curve, Hamilton used the conserved LRL vector. We can see this easily from the definition of the LRL vector in Eq. (9.45), if we rewrite it as

$$mk\frac{\mathbf{r}}{r} = (\mathbf{p}\times\mathbf{L}) - \mathbf{M} \tag{9.160}$$

and take the dot product of both sides with itself:

$$m^2k^2\frac{\mathbf{r}\cdot\mathbf{r}}{r^2} = \left[(\mathbf{p}\times\mathbf{L}) - \mathbf{M}\right]\cdot\left[(\mathbf{p}\times\mathbf{L}) - \mathbf{M}\right], \tag{9.161}$$

or

$$\begin{aligned} m^2k^2 &= (\mathbf{p}\times\mathbf{L})\cdot(\mathbf{p}\times\mathbf{L}) - 2\mathbf{M}\cdot(\mathbf{p}\times\mathbf{L}) + \mathbf{M}\cdot\mathbf{M} \\ &= (\mathbf{p}\times\mathbf{L})^2 - 2\mathbf{L}\cdot(\mathbf{M}\times\mathbf{p}) + M^2, \end{aligned} \tag{9.162}$$

where we have used the vector identity from Eq. (9.64) in the last line. The vector product in the first right-hand term of Eq. (9.162) can be written as $|\mathbf{p}|\,|\mathbf{L}|\sin\theta$, with θ the angle between **p** and **L**. Given that **p** and **L** are *orthogonal* vectors (cf. §9.3.4), we obtain

$$(\mathbf{p}\times\mathbf{L})^2 = \left(|\mathbf{p}|\,|\mathbf{L}|\sin\frac{\pi}{2}\right)^2 = (pL)^2 = p^2L^2. \tag{9.163}$$

The second term in Eq. (9.162) reads

$$\mathbf{L}\cdot(\mathbf{M}\times\mathbf{p}) = L_x(\mathbf{M}\times\mathbf{p})_x + L_y(\mathbf{M}\times\mathbf{p})_y + L_z(\mathbf{M}\times\mathbf{p})_z. \tag{9.164}$$

We know from §9.3.4 that the angular momentum vector **L** is pointing in the z-direction; that is, $\mathbf{L} = (0,0,L)$, which yields

$$\mathbf{L}\cdot(\mathbf{M}\times\mathbf{p}) = L_z(\mathbf{M}\times\mathbf{p})_z = L_z(M_xp_y - M_yp_x) = LMp_y, \tag{9.165}$$

where we have used the fact that the LRL vector $\mathbf{M} = (M,0,0)$ is directed along the x-axis. Substitution of these results in Eq. (9.162) gives

$$m^2k^2 = p^2L^2 - 2LMp_y + M^2. \tag{9.166}$$

Division by L^2 and substitution of p^2 for $p_x^2 + p_y^2$ finally yields the locus equation for $\mathbf{p} = (p_x, p_y, 0)$:

$$\frac{m^2k^2}{L^2} = p_x^2 + p_y^2 - 2\frac{M}{L}p_y + \frac{M^2}{L^2}, \tag{9.167}$$

---

[48] See Hamilton, "The Hodograph or a New Method of Expressing in Symbolic Language the Newtonian Law of Attraction." 1847. The concept of a hodograph was also used by Richard Feynman (1918–1988) in a lecture titled "The Motion of Planets Around the Sun." Delivered on March 13, 1964, to the Caltech freshman class, Feynman's "lost lecture" was later published in book form by David L. Goodstein and Judith R. Goodstein. See D. L. Goodstein and J. R. Goodstein. *Feynman's Lost Lecture: The Motions of Planets Around the Sun*. New York: W. W. Norton, 1996.

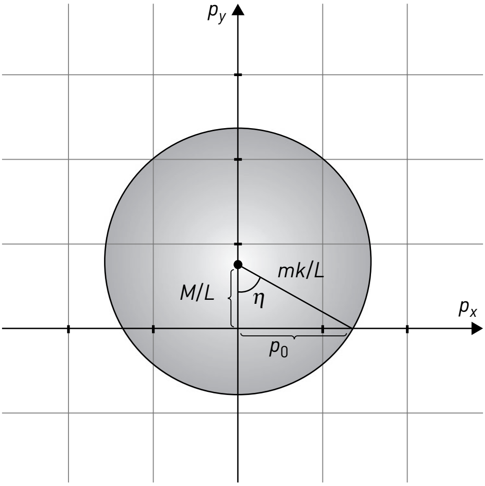


FIGURE 9.13 Depiction of a circular hodograph in momentum space, centered at the point $(0, M/L)$, with radius mk/L. The angle η is related to the eccentricity ε of the orbit via the equation $\cos\eta = \varepsilon$. The intersection of the hodograph with the p_x-axis is denoted by the points $\pm p_0 = \pm\sqrt{-2mE}$.

or

$$\left(\frac{mk}{L}\right)^2 = p_x^2 + \left(p_y - \frac{M}{L}\right)^2. \tag{9.168}$$

In other words, the momentum vector **p** traces a circle in momentum space, centered at $(0, M/L)$ on the p_y-axis and with radius mk/L (Figure 9.13). The cosine of the angle η corresponds to the eccentricity ε of the orbit as

$$\cos\eta = \frac{(M/L)}{(mk/L)} = \frac{M}{mk} = \varepsilon. \tag{9.169}$$

By changing the values of M and L, different circular hodographs can be obtained. A closer inspection of Figure 9.13 also reveals that the hodograph intersects the p_x-axis at two points at a distance denoted p_0. The magnitude of p_0 can be determined by setting $p_y = 0$ in Eq. (9.168), yielding

$$p_0^2 \equiv p_x^2 = \frac{m^2k^2}{L^2} - \frac{M^2}{L^2}, \tag{9.170}$$

and from the identity in Eq. (9.58) for $\mathbf{M}^2 = M^2$, we obtain

$$p_0^2 = \frac{m^2k^2}{L^2} - \frac{2mEL^2}{L^2} - \frac{m^2k^2}{L^2} = -2mE, \tag{9.171}$$

or

$$p_0 = \sqrt{-2mE}. \tag{9.172}$$

Clearly, all the hodographs (with differing M and L) that intersect the p_x-axis at these two points, have the same value for p_0 and thus share the same energy E. They form a *degenerate set* of hodographs (Figure 9.14). Because degeneracy is a sign of symmetry, the hodographs can be used to illustrate the symmetry of the Kepler-Coulomb problem. Worded somewhat differently, the multiple hodographs of energy E are related to one another and can be transformed into one another under a particular symmetry transformation. To find out the exact nature of this symmetry, we have to lift the hodographs into the fourth dimension via a stereographic projection. The relevant definitions and theorems are introduced in the next section.

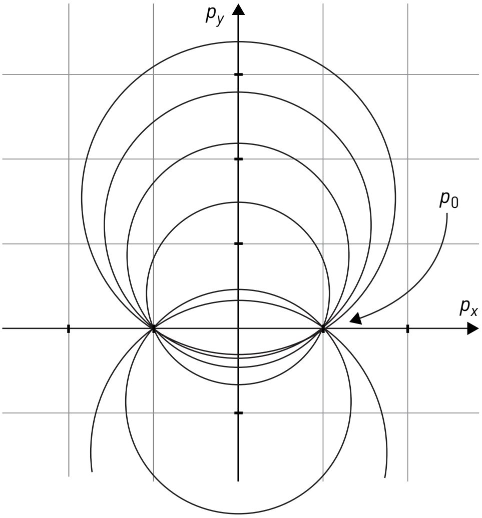


FIGURE 9.14 A family of degenerate hodographs for a given energy E. Each of the hodographs passes through the same two points $\pm p_0 = \pm\sqrt{-2mE}$ on the p_x-axis.

9.6.4 Stereographic projections in hyperspace

Hyperspheres

In Chapter 3, we started our journey with the circle, and then quickly moved on to the sphere, symbol of perfection in the three-dimensional world. But as the present chapter shows, physical phenomena can realize symmetries that are even more perfect. They lead us into higher dimensional spaces, where we meet so-called hyperspheres.

Definition 9.6 (The unit n-sphere): Let $n \in \mathbb{N}$ represent any natural number. The *n-sphere* of radius 1 $\mathbb{S}^n$ is then defined as the subset of $\mathbb{R}^{n+1}$, for which

$$\mathbb{S}^n = \left\{\mathbf{x} \in \mathbb{R}^{n+1} \mid x_1^2 + x_2^2 + \ldots + x_{n+1}^2 = 1\right\}, \tag{9.173}$$

with $\mathbf{x}$ a position vector in $(n+1)$-dimensional Euclidean space $\mathbb{R}^{n+1}$. ■

The n-sphere represents, in other words, the n-dimensional surface of an $n+1$-dimensional object. Topologically, the n-sphere corresponds to an n-dimensional *manifold*. Familiar examples include the unit circle $\mathbb{S}^1$ in two-dimensional space $\mathbb{R}^2$, and the unit sphere $\mathbb{S}^2$ in three-dimensional space $\mathbb{R}^3$. For $n > 2$, we generally speak of *hyperspheres*. In dealing with the Kepler problem, we are primarily concerned with the 3-sphere $\mathbb{S}^3$ (also known as a *glome*), which is suspended in four-dimensional Euclidean space $\mathbb{R}^4$.[49]

Great circles

Let an n-sphere $\mathbb{S}^n$ be intersected by a plane $\mathbb{II}$. Then, a circle c on the sphere is formed by the set of all points $\mathbf{x} = (x_1, x_2, \ldots, x_{n+1})$ of $\mathbb{S}^n$ that also lie on the plane $\mathbb{II}$. When the slicing plane passes through the center of the sphere, it traces a *great circle* on the

---

[49] This space should not be confused with *Minkowski space-time* (as encountered in Einstein's special theory of relativity), because this space consists of three spatial dimensions and one time dimension, and therefore has a different metric than $\mathbb{R}^4$.

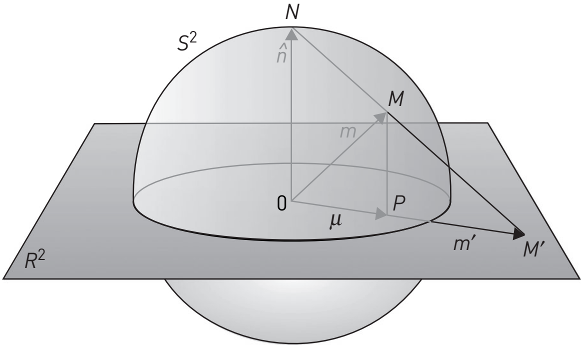


FIGURE 9.15 Stereographic projection of the sphere $\mathbb{S}^2$ onto the plane $\mathbb{R}^2$.

sphere. Examples of great circles are the equator and the meridians of longitude; the parallels of latitude, on the other hand, are *small circles* (with the exception of the equator).

Definition 9.7 (Great circles): A *great circle* (or a *Riemannian circle*) on the n-sphere $\mathbb{S}^n$ is defined as the intersection of $\mathbb{S}^n$ with a two-dimensional plane $\mathbb{\Pi}$ that passes through the center of the sphere in $(n+1)$-dimensional Euclidean space $\mathbb{R}^{n+1}$. ■

Stereographic projections

The 2-sphere $\mathbb{S}^2$, embedded in three-dimensional space, can be mapped onto a two-dimensional plane via a *stereographic projection* (think of a map representing Earth). This can be generalized to $n+1$ dimensions, where the n-sphere in $(n+1)$-dimensional space is mapped onto an n-dimensional *hyperplane* via the same stereographic principles. Let us see how this works in our familiar three dimensions (see also Figure 9.15).

Definition 9.8 (Stereographic projection in 3-space): Let $\mathbb{S}^2 = \left\{\mathbf{x} = (x,y,z) \mid x^2+y^2+z^2=1\right\}$ represent the unit sphere in 3-space and let $\hat{\mathbf{n}} = (0,0,1) \in \mathbb{R}^3$ denote the *north pole*. Construct the two-dimensional plane $\mathbb{R}^2 = \left\{(x,y,0) \in \mathbb{R}^3\right\}$, which cuts the 2-sphere in half. The intersection $\mathbb{R}^2 \cap \mathbb{S}^2$ defines the *equator* of $\mathbb{S}^2$.[50] Given any point $\mathbf{m} = (m_x, m_y, m_z) \in \mathbb{S}^2$, other than the north pole $\hat{\mathbf{n}}$, there is a unique line connecting $\mathbf{m}$ with $\hat{\mathbf{n}}$ that intersects the equatorial xy-plane $\mathbb{R}^2$ at a point $\mathbf{m}' = \left(m'_x, m'_y, 0\right)$. The *stereographic projection* of $\mathbb{S}^2$ is then defined as the map:

$$\xi : \mathbb{S}^2 \backslash \{\hat{\mathbf{n}}\} \to \mathbb{R}^2 : \mathbf{m} \mapsto \mathbf{m}', \tag{9.174}$$

where $\mathbf{m}'$ is the *projection* of $\mathbf{m}$. Here, the notation $\mathbb{S}^2 \backslash \{\hat{\mathbf{n}}\}$ denotes all points of $\mathbb{S}^2$ except $\hat{\mathbf{n}}$. Note that the map ξ is *not* defined for the projection point $\hat{\mathbf{n}}$.[51] ■

---

[50] In some cases, the unit sphere $\mathbb{S}^2$ is placed *on top* of the plane $\mathbb{R}^2$; $\mathbb{R}^2$ is then *tangent* to $\mathbb{S}^2$ at the south pole: $\hat{\mathbf{s}} = (0,0,-1)$ (see, for example, Figure 9.17), which is presented later in the chapter.

[51] Alternatively, $\hat{\mathbf{n}}$ could be said to map to *infinity* in the plane; that is, $\xi\left(\hat{\mathbf{n}}\right) = \infty$, in which case the stereographic projection is defined as the map $\xi : \mathbb{S}^2 \to \mathbb{R}^2 \cup \{\infty\} : \mathbf{m} \mapsto \mathbf{m}'$.

One way of picturing what happens during a stereographic projection is to imagine a *light source* placed at the north pole $\hat{\mathbf{n}}$. For any point $\mathbf{m} \in \mathbb{S}^2 \backslash \{\hat{\mathbf{n}}\}$, there is a *light ray* emanating downward from the north pole and passing through the sphere at $\mathbf{m}$. The ray finally meets the plane $\mathbb{R}^2$ at the point $\xi(\mathbf{m}) = \mathbf{m}'$. Note that points on the southern hemisphere are projected inside the equator.

It is also possible to work the other way around and to map the point $\mathbf{m}'$ in the plane $\mathbb{R}^2$ onto the point $\mathbf{m}$ of the unit sphere $\mathbb{S}^2$. The *inverse stereographic projection* is thus defined as the map

$$\xi^{-1} : \mathbb{R}^2 \to \mathbb{S}^2 \backslash \{\hat{\mathbf{n}}\} : \mathbf{m}' \mapsto \mathbf{m}. \tag{9.175}$$

Let us elucidate the nature of the map ξ^{-1}. To this aim, consider Figure 9.15 once again. Note that the vector $\boldsymbol{\mu} = \left(m_x, m_y, 0\right)$ is the component of $\mathbf{m}$ in the plane $\mathbb{R}^2$. The triangles $\triangle ONM'$ and $\triangle PMM'$ both have a right angle and share the angle α. They are, therefore, *similar*:

$$\triangle ONM' \sim \triangle PMM'. \tag{9.176}$$

On account of Eq. (9.176), we have

$$\frac{PM'}{OM'} = \frac{PM}{ON}. \tag{9.177}$$

Since $PM' = \left|\mathbf{m}' - \boldsymbol{\mu}\right|$, $OM' = \left|\mathbf{m}'\right|$, $PM = m_z$, and $ON = 1$, we obtain

$$\frac{\left|\mathbf{m}' - \boldsymbol{\mu}\right|}{|\mathbf{m}'|} = \frac{|m_z|}{1}. \tag{9.178}$$

Next consider the right-angle triangle $\triangle PMO$. Given that the hypothenuse starts at O and ends at M on the sphere, its length equals one. By Pythagoras,

$$1^2 = m_z^2 + |\boldsymbol{\mu}|^2 \quad \Rightarrow \quad |m_z| = \sqrt{1 - |\boldsymbol{\mu}|^2}. \tag{9.179}$$

Substitution of Eq. (9.179) in Eq. (9.178), then squaring yields

$$\frac{\left|\mathbf{m}'\right|^2 + |\boldsymbol{\mu}|^2 - 2\left|\mathbf{m}'\right||\boldsymbol{\mu}|}{|\mathbf{m}'|^2} = 1 - |\boldsymbol{\mu}|^2, \tag{9.180}$$

or

$$|\boldsymbol{\mu}|\left[|\boldsymbol{\mu}| - 2\left|\mathbf{m}'\right| + \left|\mathbf{m}'\right|^2 |\boldsymbol{\mu}|\right] = 0. \tag{9.181}$$

Since $|\boldsymbol{\mu}| > 0$, the expression in square brackets must be zero, and thus

$$|\boldsymbol{\mu}|\left[1 + \left|\mathbf{m}'\right|^2\right] = 2\left|\mathbf{m}'\right| \quad \Rightarrow \quad |\boldsymbol{\mu}| = \frac{2\left|\mathbf{m}'\right|}{1 + |\mathbf{m}'|^2}. \tag{9.182}$$

The vectors $\boldsymbol{\mu}$ and $\mathbf{m}'$ have the same direction

$$\boldsymbol{\mu} = \frac{2\mathbf{m}'}{1 + m'^2}, \tag{9.183}$$

with $\left|\mathbf{m}'\right| = m'$. Substitution of Eq. (9.182) in Eq. (9.179) then squaring gives

$$m_z^2 = 1 - \left[\frac{2\left|\mathbf{m}'\right|}{1 + |\mathbf{m}'|^2}\right]^2 = \frac{1 + \left|\mathbf{m}'\right|^4 + 2\left|\mathbf{m}'\right|^2 - 4\left|\mathbf{m}'\right|^2}{\left[1 + |\mathbf{m}'|^2\right]^2} = \left[\frac{\left|\mathbf{m}'\right|^2 - 1}{|\mathbf{m}'|^2 + 1}\right]^2, \tag{9.184}$$

and thus

$$m_z = \frac{m'^2 - 1}{m'^2 + 1}. \tag{9.185}$$

On account of Eqs. (9.183) and (9.185), the map $\xi^{-1}: \mathbb{R}^2 \to \mathbb{S}^2 \backslash \{\hat{\mathbf{n}}\} : \mathbf{m}' \mapsto \mathbf{m}$ is given by

$$m_x = \frac{2m'_x}{1+m'^2}; \quad m_y = \frac{2m'_y}{1+m'^2}; \quad m_z = \frac{m'^2-1}{1+m'^2}. \tag{9.186}$$

In general

$$x = \frac{2x'}{1+x'^2+y'^2}; \quad y = \frac{2y'}{1+x'^2+y'^2}; \quad z = \frac{x'^2+y'^2-1}{1+x'^2+y'^2}. \tag{9.187}$$

Do not disturb my circles!

One particularly interesting property of stereographic projections, which we will need later, is the following.

Theorem 9.1 *Great circles* on the unit sphere $\mathbb{S}^2$ are mapped onto *circles* in the plane $\mathbb{R}^2$ (and vice versa) under the stereographic projection ξ. ■

Proof. 1. Plane equation: Recall from Definition 9.7 that a circle c on $\mathbb{S}^2$ is defined as the intersection $\mathbb{S}^2 \cap \mathbb{II}$ of the sphere with a slicing plane $\mathbb{II}$. The equation of the plane can be determined as follows. Let $\mathbf{N} = \left(N_x, N_y, N_z\right)$ denote a *normal vector* of the plane $\mathbb{II}$ (i.e., $\mathbf{N}$ is perpendicular to $\mathbb{II}$), which goes through the point $\mathbf{x}_0 = \left(x_0, y_0, z_0\right)$ of $\mathbb{II}$. In addition, let $\mathbf{x} = \left(x, y, z\right)$ be the position vector of any other point on the plane $\mathbb{II}$. Then, $\mathbf{x} - \mathbf{x}_0$ is a vector that connects both points and lies in the plane $\mathbb{II}$:

$$\mathbf{x} - \mathbf{x_0} = \left(x - x_0, y - y_0, z - z_0\right). \tag{9.188}$$

Since $\mathbf{x} - \mathbf{x}_0$ and the normal vector $\mathbf{N}$ are perpendicular to each other, their dot product should be zero

$$\mathbf{N} \cdot (\mathbf{x} - \mathbf{x}_0) = 0. \tag{9.189}$$

Expanding gives

$$N_x (x - x_0) + N_y \left(y - y_0\right) + N_z (z - z_0) = 0, \tag{9.190}$$

or

$$N_x x + N_y y + N_z z - \left(N_x x_0 + N_y y_0 + N_z z_0\right) = 0. \tag{9.191}$$

Substitution of the constant term $-\left(N_x x_0 + N_y y_0 + N_z z_0\right)$ for N_0 finally yields the equation for the plane $\mathbb{II}$:

$$N_x x + N_y y + N_z z + N_0 = 0. \tag{9.192}$$

2. Circle equation: A circle in the *xy*-plane with radius ρ and center $\left(x_0, y_0\right)$ is defined by the equation

$$(x - x_0)^2 + \left(y - y_0\right)^2 = \rho^2. \tag{9.193}$$

3. Stereographic projection: To obtain the equation for the projection of a circle c onto the *xy*-plane under a stereographic projection, we substitute Eq. (9.187) into Eq. (9.192), yielding

$$N_x \frac{2x'}{1+x'^2+y'^2} + N_y \frac{2y'}{1+x'^2+y'^2} + N_z \frac{x'^2+y'^2-1}{1+x'^2+y'^2} + N_0 = 0. \tag{9.194}$$

Multiplying through by $1 + x'^2 + y'^2$ gives

$$N_x 2x' + N_y 2y' + N_z \left(x'^2 + y'^2 - 1\right) + N_0 \left(x'^2 + y'^2 + 1\right) = 0, \tag{9.195}$$

or

$$N_x 2x' + N_y 2y' + \left(x'^2 + y'^2\right)(N_z + N_0) = N_z - N_0. \tag{9.196}$$

Dividing by $N_z + N_0$ gives us

$$x'^2 + y'^2 + \frac{2N_x}{N_z + N_0} x' + \frac{2N_y}{N_z + N_0} y' = \frac{N_z - N_0}{N_z + N_0}. \tag{9.197}$$

Completing the square finally yields

$$\left(x' + \frac{N_x}{N_z + N_0}\right)^2 + \left(y' + \frac{N_y}{N_z + N_0}\right)^2 = \frac{N_x^2 + N_y^2 + N_z^2 - N_0^2}{(N_z + N_0)^2}. \tag{9.198}$$

On comparison with Eq. (9.193), we notice that this is the equation for a circle in the *xy*-plane with radius $\rho = \sqrt{(N_x^2 + N_y^2 + N_z^2 - N_0^2) \big/ (N_z + N_0)^2}$, centered at $\big(-N_x/(N_z + N_0), -N_y/(N_z + N_0)\big)$ This completes our proof that circles on $\mathbb{S}^2$ map onto circles in the plane. Reversibly, circles in $\mathbb{R}^2$ project onto circles of $\mathbb{S}^2$, because the mappings ξ and ξ^{-1} are *one* to *one* and *continuous*. ▪

The Kepler problem in four dimensions

To find the symmetry transformations that connect the degenerate hodographs to one another, we first need to introduce an *extended momentum space* in four dimensions $\mathbb{R}^4$ by taking the three momentum vector components p_1, p_2, and p_3, and adding a fourth component p_4 to it. A hodograph in momentum 3-space $\mathbb{R}^3$ can then be represented within momentum 4-space $\mathbb{R}^4$ via an *inverse stereographic projection* of the hodograph onto a 3-sphere $\mathbb{S}^3$. It is of crucial importance to project a hodograph with given p_0 onto a hypersphere with radius equal to p_0.

Definition 9.9 (Inverse stereographic projection in momentum 4-space): Let $\mathbb{S}^3 = \left\{\mathbf{p} = (p_1, p_2, p_3, p_4) \mid p_1^2 + p_2^2 + p_3^2 + p_4^2 = p_0^2\right\}$ represent a hypersphere in momentum 4-space with radius $p_0 = \sqrt{-2mE}$ and let $\hat{\mathbf{n}} = (0, 0, 0, p_0)$ denote the *north pole*. Construct the three-dimensional hyperplane $\mathbb{R}^3 = \left\{(p_1, p_2, p_3, 0) \in \mathbb{R}^4\right\}$, which cuts the 3-sphere in half. The intersection $\mathbb{R}^3 \cap \mathbb{S}^3$ defines the *equator* of $\mathbb{S}^3$. Given any point $\mathbf{p} = (p_1, p_2, p_3, p_4) \in \mathbb{S}^3$, other than the north pole $\hat{\mathbf{n}}$, there is a unique line connecting $\mathbf{p}$ with $\hat{\mathbf{n}}$ that intersects the equatorial hyperplane $\mathbb{R}^3$ at a point $\mathbf{p}' = (p_1', p_2', p_3', 0)$. The *inverse stereographic projection* of $\mathbb{S}^3$ is then defined as the map

$$\xi^{-1} : \mathbb{R}^3 \to \mathbb{S}^3 \backslash \{\hat{\mathbf{n}}\} : \mathbf{p}' \mapsto \mathbf{p}, \tag{9.199}$$

where $\mathbf{p}$ is the *inverse projection* of $\mathbf{p}'$. ▪

Fortunately, the hodographs for the Kepler problem were seen to be circular (§9.6.3). On account of Theorem 9.1, we expect the hodographs in momentum 3-space $\mathbb{R}^3$ to be sent into *circles* of the hypersphere $\mathbb{S}^3$ under the inverse stereographic projection ξ^{-1}. To be more explicit, the hodographs project onto *great circles* of the hypersphere.[52]

---

[52] Notice that a free particle, constrained to the surface of a hypersphere $\mathbb{S}^3$, traces *great circles* as classical trajectories. This explains Fock's assertion that the motion of a particle bound in a Coulomb field in flat three-dimensional space $\mathbb{R}^3$ is mathematically equivalent to the motion of a free particle on a hypersphere $\mathbb{S}^3$.

It follows that a degenerate set of hodographs (with the same value for p_0) corresponds to a set of great circles on the 3-sphere $\mathbb{S}^3$ of radius p_0. The great circles can be simply rotated into one another via a rigid rotation of the hypersphere about its origin. This, in turn, when reprojected stereographically onto the three-dimensional hyperplane $\mathbb{R}^3$, manifests itself by a continuous mapping of the degenerate hodographs onto one another, while keeping the energy E invariant. Schematically,[53]

$$\text{Circle in } \mathbb{R}^3 \xrightarrow{\xi^{-1}} \text{Circle in } \mathbb{S}^3 \xrightarrow{\mathrm{SO}(4)} \text{Circle in } \mathbb{S}^3 \xrightarrow{\xi} \text{Circle in } \mathbb{R}^3. \qquad (9.200)$$

We have thus identified the hidden symmetry of the Kepler problem as the rotation group in four dimensions SO(4).

9.6.5 Peeking at the fourth dimension

Quantum mechanically, the mapping in Eq. (9.200) of the hodographs onto one another under the SO(4) group corresponds to a mixing of the degenerate orbitals belonging to the same n-level. Let us study this important point by way of the following example.

Consider the set of hydrogenic orbitals with $n = 2$: $2s$, $2p_x$, $2p_y$, and $2p_z$. Since the energy of these four orbitals is the same, they are said to be *degenerate*, and Noether's theorem informs us that there should be a symmetry that relates these orbitals by transforming them into one another. Another way of stating this is by saying that these symmetry-related orbitals are *equivalent*; they simply represent two sides of the same coin.

However, an s-orbital is spherical and a p-orbital has two lobes; so, although it is easy to see that one p-orbital can be rotated into another p-orbital under the symmetry operations of the SO(3) group, it is far from obvious how a p-orbital can be "rotated" into an s-orbital. Although the $(2l+1)$-fold degeneracy is a logical consequence of the rotational symmetry of the Coulomb potential, the n^2-degeneracy seems to be completely *accidental*. We are thus forced to postulate the existence of an additional symmetry, *hidden* from direct observation, that relates the s-orbitals to the p-orbitals.

In our quest for this dynamical symmetry, we discovered that the Coulomb potential has rotational symmetry not only in three dimensions, but also in *four*! Indeed, when viewed from the perspective of the SO(4) group, all the atomic orbitals of fixed n are grouped together in one multiplet of dimension n^2. This means that there are rotations in four dimensions that transform the different varieties of orbitals into one another, thus explaining the n^2-degeneracy.[54]

Imagining the fourth dimension is not an easy task, however. Many artists and writers alike have attempted to picture the fourth dimension in one way or another. In 1884, a schoolmaster in the city of London whose name was Edwin Abbott Abbott (1838–1926) published a satirical novella titled *Flatland: A Romance of Many Dimensions*.[55] It recounted the story of Arthur Square, an ordinary square who lived

---

[53] See R. Gilmore. *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists*. Cambridge: Cambridge University Press, 2008, p. 265.

[54] See S. M. Blinder. "Quantum Alchemy: Transmutation of Atomic Orbitals." *Journal of Chemical Education* 78.3 (2001), pp. 391–394.

[55] E. A. Abbott. *Flatland: A Romance of Many Dimensions*, 1884.

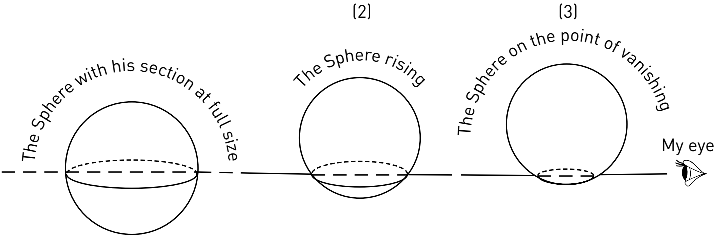


FIGURE 9.16 A three-dimensional sphere intersecting Flatland's two-dimensional world. Illustration from E. A. Abbott. *Flatland: A Romance of Many Dimensions.* Seely & Co., 1884.

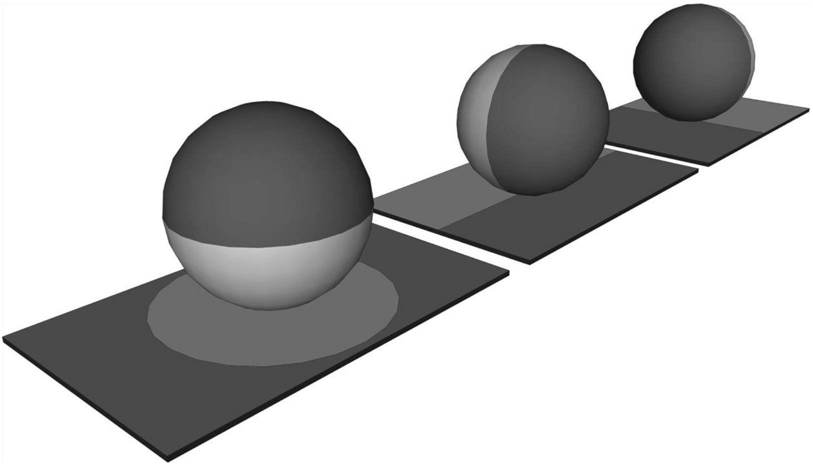


FIGURE 9.17 An illustration of how it is possible to "rotate" *s*- and *p*-orbitals into one another by stepping up a dimension. The orbitals are pictured by different patterns on the two-dimensional plane. The circular pattern represents an *s*-orbital, and the remaining two patterns, two *p*-orbitals. Acknowledging that these patterns result from projections of a three-dimensional sphere onto the two-dimensional space, we can see how the patterns in two dimensions can be interchanged by rotating the sphere. The Kepler-Coulomb potential has four-dimensional symmetry and permits this type of rotation. Adapted from P. Atkins. *Galileo's Finger: The Ten Great Ideas of Science.* Oxford: Oxford University Press, 2003, pp. 173–175.

in a two-dimensional world called *Flatland*. One day, Arthur was visited by a sphere from the third dimension—a magical creature with mind-blowing, almost God-like powers. The sphere could talk to Arthur while remaining invisible by hovering above Flatland. When traveling through Flatland, the sphere seemed to appear from nowhere, growing in size and then shrinking again, finally to disappear (Figure 9.16). At one point, the sphere bumped into Arthur and catapulted him into the third dimension: *Spaceland*. Being introduced in this way to the third dimension, Arthur used the idea of *dimensional analogy* to imagine what the fourth dimension would be like. We can use a similar analogy to picture what happens when an *s*-orbital is "rotated" into a *p*-orbital.

Imagine a sphere resting on a two-dimensional surface (Figure 9.17). In ordinary 3-space, the sphere could represent Earth, which is projected stereographically onto the plane to obtain a two-dimensional map of the world, in much the same way as Flemish cartographer Gerardus Mercator (1512–1594) compiled the first modern atlas. Now imagine stepping up one dimension. The *plane* now represents our familiar *three-dimensional* world, which is assumed to be a *projection* of an even higher *four-dimensional* world, as represented by the sphere. Notice also that the northern hemisphere is dark and the southern hemisphere is light. When projecting the surface of the sphere onto the plane, one obtains a light circle centered on a dark background. Imagine next rotating the sphere through 90° to obtain the sphere in the second part of the illustration. A projection of this sphere bisects the plane in two halves: one light and another dark. A further rotation of the sphere through another 90° yields the third part of Figure 9.17, which has a similar projection as the second sphere, although the map has been rotated by 90°.

Now let the first orientation of the sphere represent an *s*-orbital, and the remaining two, two different *p*-orbitals—say, a p_x-and a p_y-orbital. We Flatlanders, who are stuck in the three-dimensional plane, have no difficulty in seeing that the two *p*-orbitals are related to one another. After all, a mere rotation of the plane transforms both orbitals into one another. We are baffled, however, by the thought that the *s*-orbital, with its circular shape, could be "rotated" somehow into one of the two *p*-orbitals. We have great difficulties, in other words, understanding that the *s*-orbital has the same energy as the two *p*-orbitals. Superhumans living in the four dimensions do not face these troubles. They see all our Flatland patterns as projections of a sphere, which are related by simple rotations in the fourth dimension. As Peter Atkins observed: "the equality of energies of apparently unrelated orbitals is a consequence of there being symmetry hidden away in a fourth dimension."[56]

---

[56] P. Atkins. *Galileo's Finger: The Ten Great Ideas of Science*, p. 175.

10 Scholium II

"There are in my opinion three possibilities . . .:

1. *Brinton Jr. wants to harm his father, and everything he has told me is nothing but comedy.*
2. *Seaburry's amiability is only feigned, he is in reality a satanic evildoer.*
3. *Brinton Sr. is only imagining things. Admittedly, the world-renowned Dundalk has rejected this possibility, but why couldn't he be wrong for once?"*

"There are, my dear Watson, not three possibilities, but two *key questions. The first: with whose mouth does Shiva speak? The second: why does Shiva speak?"*
–Clifford Semper[1]

The treatment of the hydrogen atom has yielded a connection between orthogonal and unitary groups. In the first part of this scholium, our detective work will trace the geometric form of this connection, making use of a transformation introduced in 1964 by Finnish astronomer and mathematician Paul Kustaanheimo (1924–1997) and further developed in collaboration with Swiss mathematician Eduard Stiefel (1909–1978). This again offers the possibility to witness a symmetry shattering, this time from SU(4) to SU(2) $\otimes$ SU(2).

---

[1] Translated from C. Semper. "De Goden van Brinton Hall." In: *De Nalatenschap van Sherlock Holmes*. Amsterdam: Andries Blitz.

10.1 OSCILLATOR ROOTS OF THE HYDROGEN PROBLEM

In quantum chemistry the hydrogen atom and the harmonic oscillator are usually represented as two separate problems. However, the factorization of the $\mathfrak{so}(4)$ Lie algebra of the hydrogen Hamiltonian into two separate $\mathfrak{su}(2)$ oscillator algebras indicates the existence of a deep connection between the two problems. As we shall see in this scholium, an explicit construction of this connection can be achieved by means of the so-called *Kustaanheimo-Stiefel* (KS) transformation, which was originally developed in the context of celestial mechanics. Although in the Fock analysis the momentum space was extended to a four-dimensional space, where the Kepler problem appears as a projection, the KS transformation acts in coordinate space and resolves it into an underlying set of spinors, which also form a four-dimensional space. In both cases, a constraint appears that implies the sacrifice of the extra dimension. As a preliminary, let us first explore the possible correspondence between orthogonal and unitary Lie algebras from a purely dimensional point of view.

10.1.1 Dimensional considerations

As indicated earlier, the orders of the respective Lie algebras are given by

$$\begin{aligned} \mathfrak{so}(n) &: \frac{n(n-1)}{2}; \\ \mathfrak{su}(m) &: m^2-1. \end{aligned} \tag{10.1}$$

For both algebras to be isomorphic to each other, the minimal requirement is that the two orders are equal:

$$m = \sqrt{\frac{n(n-1)}{2}+1}. \tag{10.2}$$

This is a so-called *Diophantine equation*; in other words, for a given integer n, only integer solutions for m are allowed. Table 10.1 contains a list of the first few integers for which this equation holds.

In Chapter 8, we saw the equivalence $\mathfrak{so}(3) \sim \mathfrak{su}(2)$. Interestingly, as Table 10.1 indicates, there is a further homomorphism between $\mathfrak{so}(6)$ and $\mathfrak{su}(4)$,[2] but for the

Table 10.1 Special orthogonal and unitary algebra's of the same order

$\mathfrak{so}(n)$	$\mathfrak{su}(m)$	m^2-1	*Homomorphism*
$\mathfrak{so}(3)$	$\mathfrak{su}(2)$	3	Yes
$\mathfrak{so}(6)$	$\mathfrak{su}(4)$	15	Yes
$\mathfrak{so}(16)$	$\mathfrak{su}(11)$	120	No
$\mathfrak{so}(33)$	$\mathfrak{su}(23)$	528	No
$\mathfrak{so}(91)$	$\mathfrak{su}(64)$	4095	No

---

[2] In Chapter 12, we will make use of the SO(6) symmetry in its dynamical form as SO(4,2), which is isomorphic to SU(2,2).

other cases in the table the Lie algebras are *not* equivalent. Hence, the connection between rotational and unitary symmetries is limited to special cases.[3]

10.1.2 The SU(4) oscillator

As mentioned, the SO(4) symmetry group is structurally equivalent to the product of two SU(2) oscillator groups. The four-dimensional oscillator, SU(4), constitutes a convenient covering group to embed these two SU(2) oscillators. The corresponding Hamiltonian equation reads

$$\begin{aligned} \frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial u_1^2}+\frac{\partial^2}{\partial u_2^2}+\frac{\partial^2}{\partial u_3^2}+\frac{\partial^2}{\partial u_4^2}\right)|\psi\rangle+\frac{1}{2}m\omega^2\left(u_1^2+u_2^2+u_3^2+u_4^2\right)|\psi\rangle \\ =\hbar\omega\sum_{i=1}^{4}\left(n_i+\frac{1}{2}\right)|\psi\rangle. \end{aligned} \tag{10.3}$$

Here, the u_i coordinates are components of a four-dimensional vector. Each component contributes to the total energy by $\hbar\omega(n_i+1/2)$, where n_i is the excitation quantum number along this component. Let us now group the components together in two subsets (u_1,u_2) and (u_3,u_4), and rewrite them in cylindrical coordinate form as

$$\begin{aligned} u_1 &= u\cos\beta; \\ u_2 &= u\sin\beta; \end{aligned} \tag{10.4}$$

$$\begin{aligned} u_3 &= v\cos\gamma; \\ u_4 &= v\sin\gamma. \end{aligned} \tag{10.5}$$

The Hamiltonian can be easily expressed in these cylindrical coordinates. According to the chain rule, we have for the derivative operators

$$\begin{aligned} \frac{\partial}{\partial u_1} &= \frac{\partial u}{\partial u_1}\frac{\partial}{\partial u}+\frac{\partial\beta}{\partial u_1}\frac{\partial}{\partial\beta}; \\ \frac{\partial}{\partial u_2} &= \frac{\partial u}{\partial u_2}\frac{\partial}{\partial u}+\frac{\partial\beta}{\partial u_2}\frac{\partial}{\partial\beta}, \end{aligned} \tag{10.6}$$

and similarly for u_3 and u_4. To obtain the required partial derivatives, we need the inverse transformations:

$$\begin{aligned} u^2 &= u_1^2+u_2^2; \\ \tan\beta &= \frac{u_2}{u_1}. \end{aligned} \tag{10.7}$$

By differentiating, we obtain

$$\begin{aligned} \frac{\partial}{\partial u_1}u^2 &= 2u\frac{\partial u}{\partial u_1} = 2u_1; \\ \frac{\partial}{\partial u_2}u^2 &= 2u\frac{\partial u}{\partial u_2} = 2u_2; \end{aligned} \tag{10.8}$$

---

[3] It is striking that for three cases in Table 10.1, with $m=2,4,64$, the dimension of the oscillator algebra is a power of two. A century ago, in 1913, the mathematical genius Srinivasa Ramanujan, with his gift for number patterns, conjectured that only a few cases with $m=2^p$ would exist. In 1948, Trygve Nagell proved that, indeed, in addition to the trivial case with $p=0$ and the equivalence $\mathfrak{so}(2)\sim\mathfrak{u}(1)$, the only allowed values of p correspond to $p=1$, 2, and 6. See Y. Pavlyukh and A. R. P. Rau. “1-, 2-, and 6-Qubits, and the Ramanujan–Nagell Theorem.” *International Journal of Quantum Information* 11.06 (2013), article no. 1350056.

$$\begin{aligned}\frac{\partial}{\partial u_1}\tan\beta &= \frac{1}{\cos^2\beta}\frac{\partial\beta}{\partial u_1} = -\frac{u_2}{u_1^2};\\ \frac{\partial}{\partial u_2}\tan\beta &= \frac{1}{\cos^2\beta}\frac{\partial\beta}{\partial u_2} = \frac{1}{u_1}.\end{aligned}\tag{10.9}$$

The derivative operators then become,

$$\begin{aligned}\frac{\partial}{\partial u_1} &= \cos\beta\frac{\partial}{\partial u} - \frac{\sin\beta}{u}\frac{\partial}{\partial\beta},\\ \frac{\partial}{\partial u_2} &= \sin\beta\frac{\partial}{\partial u} + \frac{\cos\beta}{u}\frac{\partial}{\partial\beta},\end{aligned}\tag{10.10}$$

and entirely similar for u_3 and u_4. The cylindrical form of the Laplacian is then easily obtained:

$$\begin{aligned}\sum_{i=1}^{4}\frac{\partial^2}{\partial u_i^2} &= \frac{\partial^2}{\partial u^2} + \frac{1}{u}\frac{\partial}{\partial u} + \frac{1}{u^2}\frac{\partial^2}{\partial\beta^2}\\ &\quad + \frac{\partial^2}{\partial v^2} + \frac{1}{v}\frac{\partial}{\partial v} + \frac{1}{v^2}\frac{\partial^2}{\partial\gamma^2}.\end{aligned}\tag{10.11}$$

10.1.3 The Kustaanheimo-Stiefel transformation

The Kustaanheimo-Stiefel (KS) transformation provides a connection between the four-dimensional vector (u_1, u_2, u_3, u_4) and the three-dimensional vector (x, y, z). Our treatment follows the geometric algebraic approach as outlined by Bartsch.[4] The transformation is based on the rotation of a spinor. Consider the basis spinor, represented by the Pauli matrix σ_z:

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.\tag{10.12}$$

A rotation of this spinor to an arbitrary direction in space, corresponding to the Cartesian coordinates x, y, and z, can be achieved by a unitary matrix transformation:

$$x\sigma_x + y\sigma_y + z\sigma_z = \mathbb{U}\sigma_z\mathbb{U}^\dagger,\tag{10.13}$$

where the unitary matrix is expressed in the usual Cayley-Klein parameters:

$$\begin{aligned}\mathbb{U} &= \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix};\\ \mathbb{U}^\dagger &= \begin{bmatrix} a^* & -b \\ b^* & a \end{bmatrix}.\end{aligned}\tag{10.14}$$

The matrix product is thus given by

$$\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a^* & -b \\ b^* & a \end{bmatrix} = \begin{bmatrix} |a|^2 - |b|^2 & -2ab \\ -2a^*b^* & -|a|^2 + |b|^2 \end{bmatrix}.\tag{10.15}$$

This result can be expressed in Pauli matrices as

$$\mathbb{U}\sigma_z\mathbb{U}^\dagger = \left(-ab - (ab)^*\right)\sigma_x - i\left(ab - (ab)^*\right)\sigma_y + \left(|a|^2 - |b|^2\right)\sigma_z.\tag{10.16}$$

---

[4] T. Bartsch. “The Kustaanheimo-Stiefel Transformation in Geometric Algebra.” *Journal of Physics A: Math. Gen.* 36 (2003), pp. 6963–6978.

Note that, because σ_z is traceless, the transformed result is traceless as well, so that the unit matrix σ_0 does not appear. From this result we immediately obtain expressions for the Cartesian components:

$$\begin{aligned} x &= -ab - (ab)^*; \\ y &= -i\left(ab - (ab)^*\right); \\ z &= |a|^2 - |b|^2. \end{aligned} \tag{10.17}$$

Here, we can recognize immediately the final column of the $\mathbb{O}(U)$ matrix in Chapter 8, Eq. (8.29), which indeed corresponds to the transformation of the z-component.

Because the Cayley-Klein parameters are, in principle, complex quantities, the expressions involve four real parameters that are identified with the $\{u_i\}_{i=1,4}$ components of a 4-vector.[5] In this way, the three Cartesian components are connected to a set of four coordinates. For a rotation, these coordinates are normalized to unity, but in the KS transformation this condition is relaxed so that the four coordinates can be identified with the four degrees of freedom of the SU(4) oscillator. This means that the transformation, in fact, combines a standard rotation with a dilatation. The conventional form of the KS transformation is obtained by defining the real and imaginary parts of the Cayley-Klein parameters as follows:

$$\begin{aligned} a &= \frac{1}{\sqrt{2}}(-u_1 + iu_2); \\ b &= \frac{1}{\sqrt{2}}(u_3 - iu_4). \end{aligned} \tag{10.18}$$

The resulting KS transformation thus reads

$$\begin{aligned} x &= u_1 u_3 - u_2 u_4; \\ y &= u_1 u_4 + u_2 u_3; \\ z &= \frac{1}{2}\left(u_1^2 + u_2^2 - u_3^2 - u_4^2\right). \end{aligned} \tag{10.19}$$

A simple interpretation of the KS transformation views the rotation of the spinor as a telescope slew (Figure 10.1). The σ_z spinor represents the telescope orientation toward the zenith. To view a particular star at angular direction (θ,ϕ), we have to slew the telescope to this view line. One way to perform this is by using a rotation axis halfway along the great circle joining the zenith to the star direction, and rotating around this axis over an angle of π. In standard SO(3) axis-angle notation, this rotation axis is parametrized as follows:

$$\begin{aligned} n_x &= \sin\frac{1}{2}\theta\cos\phi; \\ n_y &= \sin\frac{1}{2}\theta\sin\phi; \\ n_z &= \cos\frac{1}{2}\theta. \end{aligned} \tag{10.20}$$

According to Eq. (8.34), the corresponding Cayley-Klein parameters are given by

$$\begin{aligned} a &= -in_z; \\ b &= -n_y - in_x. \end{aligned} \tag{10.21}$$

---

[5] The four components of the rotation matrix delineate the four quaternions, which were conceived by Sir William Rowan Hamilton on October 16, 1843, as he walked by Broome Bridge near Dublin.

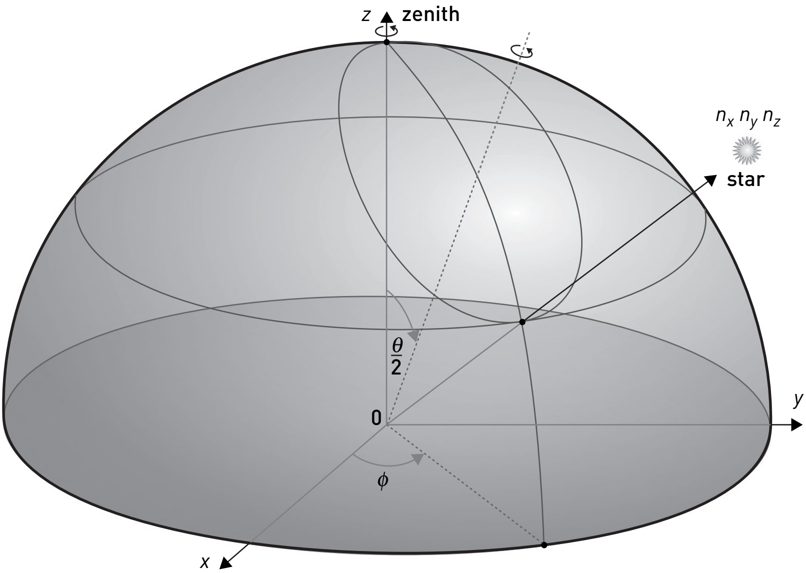


FIGURE 10.1 Telescope slew from the zenith to a star at angular position (θ, ϕ) can be performed by a rotation over 180° about the halfway axis at $\theta/2$.

This corresponds to the oscillator coordinates:

$$\begin{aligned}
u_1 &= -\sqrt{2}\,\Re(a) = 0; \\
u_2 &= \sqrt{2}\,\Im(a) = -\sqrt{2}\cos\frac{1}{2}\theta; \\
u_3 &= \sqrt{2}\,\Re(b) = -\sqrt{2}\sin\frac{1}{2}\theta\,\sin\phi; \\
u_4 &= -\sqrt{2}\,\Im(b) = \sqrt{2}\sin\frac{1}{2}\theta\,\cos\phi.
\end{aligned} \tag{10.22}$$

The corresponding Cartesian coordinates are then given by the KS transformation:

$$\begin{aligned}
x &= 2\cos\frac{1}{2}\theta\,\sin\frac{1}{2}\theta\,\cos\phi = \sin\theta\cos\phi; \\
y &= 2\cos\frac{1}{2}\theta\,\sin\frac{1}{2}\theta\,\sin\phi = \sin\theta\sin\phi; \\
z &= \cos^2\left(\frac{1}{2}\theta\right) - \sin^2\left(\frac{1}{2}\theta\right)\cos^2\phi - \sin^2\left(\frac{1}{2}\theta\right)\sin^2\phi = \cos\theta.
\end{aligned} \tag{10.23}$$

These are, of course, precisely the Cartesian coordinates corresponding to the angular position of the star. The analogy also clarifies the inherent redundancy of the KS transformation, which is inevitable in view of the reduction of a 4-vector to a 3-vector. Indeed the slew operation is not uniquely defined by the position of the star. As an example the rotation axis could also be taken perpendicular to the great circle joining the zenith to the star orientation, in which case the rotation angle would be θ instead of π. This would alter the u_i coordinates, but it does not affect the Cartesian coordinates. In the next section, this angular freedom will be shown to provide a convenient condition to connect the hydrogen atom to the four-dimensional oscillator.

10.2 THE TRANSFORMATION OF THE HYDROGEN HAMILTONIAN

Let us now resume the standard Schrödinger equation for the hydrogen atom, applied to a wave function with principal quantum number n:

$$\frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)|\psi_n\rangle-\frac{Ze^2}{4\pi\varepsilon_0 r}|\psi_n\rangle=-\frac{Z^2me^4}{8h^2{\varepsilon_0}^2n^2}|\psi_n\rangle. \tag{10.24}$$

Subsequently, we apply the KS transformation to the Cartesian coordinates, writing the u_i coordinates in polar form:

$$\begin{aligned} z&=\frac{1}{2}\left(u^2-v^2\right);\\ x&=uv\cos(\beta+\gamma);\\ y&=uv\sin(\beta+\gamma);\\ r&=\sqrt{x^2+y^2+z^2}=\frac{1}{2}\left(u^2+v^2\right). \end{aligned} \tag{10.25}$$

These coordinates are sometimes denoted as the *squared parabolic coordinates*. Note that the angles of the u and v oscillator planes only appear as a sum. This sum corresponds to the spherical coordinate angle ϕ in the x, y-plane:

$$\phi=\beta+\gamma. \tag{10.26}$$

In this way, one degree of freedom is removed. We define this degree as the difference angle and denote it as χ:

$$\chi=\beta-\gamma. \tag{10.27}$$

As we will see later, this angle plays an important role because it introduces a constraint when mapping the hydrogen atom on the four-dimensional oscillator.

To transform the Hamiltonian to parabolic coordinates, we need inverse transformations for the calculation of the required partial derivatives. We have

$$\begin{aligned} u^2&=z+r;\\ v^2&=-z+r;\\ \phi&=\arctan\frac{y}{x}. \end{aligned} \tag{10.28}$$

From these expressions, the first-order partial derivatives are easily obtained by applying the chain rule:

$$\begin{aligned} \frac{\partial}{\partial z}&=\frac{u}{2r}\frac{\partial}{\partial u}-\frac{v}{2r}\frac{\partial}{\partial v};\\ \frac{\partial}{\partial x}&=\frac{v\cos\phi}{2r}\frac{\partial}{\partial u}+\frac{u\cos\phi}{2r}\frac{\partial}{\partial v}-\frac{\sin\phi}{uv}\frac{\partial}{\partial\phi};\\ \frac{\partial}{\partial y}&=\frac{v\sin\phi}{2r}\frac{\partial}{\partial u}+\frac{u\sin\phi}{2r}\frac{\partial}{\partial v}+\frac{\cos\phi}{uv}\frac{\partial}{\partial\phi}. \end{aligned} \tag{10.29}$$

When squaring these expressions, we must, of course, take into account that the factors preceding the partial derivatives are functions of the parabolic coordinates and thus must be derived as well. We finally obtain for the Laplacian:

$$\Delta^2=\frac{1}{2r}\left(\frac{\partial^2}{\partial u^2}+\frac{1}{u}\frac{\partial}{\partial u}+\frac{\partial^2}{\partial v^2}+\frac{1}{v}\frac{\partial}{\partial v}\right)+\frac{1}{u^2v^2}\frac{\partial^2}{\partial\phi^2}. \tag{10.30}$$

The ϕ part of this expression corresponds to the unique ϕ-dependent term of the hydrogen Hamiltonian in spherical coordinates, which reads

$$\frac{1}{r^2 \sin^2\theta} \frac{\partial^2}{\partial\phi^2}. \tag{10.31}$$

This term may be rewritten as

$$\frac{1}{u^2 v^2} \frac{\partial^2}{\partial\phi^2} = \frac{1}{2r}\left(\frac{1}{u^2}\frac{\partial^2}{\partial\phi^2} + \frac{1}{v^2}\frac{\partial^2}{\partial\phi^2}\right). \tag{10.32}$$

The Laplacian thus becomes

$$\Delta^2 = \frac{1}{2r}\left(\frac{\partial^2}{\partial u^2} + \frac{1}{u}\frac{\partial}{\partial u} + \frac{\partial^2}{\partial v^2} + \frac{1}{v}\frac{\partial}{\partial v} + \frac{1}{u^2}\frac{\partial^2}{\partial\phi^2} + \frac{1}{v^2}\frac{\partial^2}{\partial\phi^2}\right). \tag{10.33}$$

When comparing this result with the Laplacian for the general SU(4) oscillator in Eq. (10.11), we note, apart from the overall division by $2r$, a difference for the angular parts, with the second derivatives in the oscillator angles β and γ being replaced by a uniform second derivative in ϕ. The relationship between these angular derivatives may be expressed via partial differentiations of the expressions for the sum and difference angles:

$$\begin{aligned}\frac{\partial}{\partial\beta} &= \frac{\partial\phi}{\partial\beta}\frac{\partial}{\partial\phi} + \frac{\partial\chi}{\partial\beta}\frac{\partial}{\partial\chi}\\ &= \frac{\partial}{\partial\phi} + \frac{\partial}{\partial\chi};\\ \frac{\partial}{\partial\gamma} &= \frac{\partial}{\partial\phi} - \frac{\partial}{\partial\chi}.\end{aligned} \tag{10.34}$$

Accordingly, the second-order derivatives in the angles of the oscillator planes become

$$\begin{aligned}\frac{\partial^2}{\partial\beta^2} &= \frac{\partial^2}{\partial\phi^2} + 2\frac{\partial^2}{\partial\phi\partial\chi} + \frac{\partial^2}{\partial\chi^2};\\ \frac{\partial^2}{\partial\gamma^2} &= \frac{\partial^2}{\partial\phi^2} - 2\frac{\partial^2}{\partial\phi\partial\chi} + \frac{\partial^2}{\partial\chi^2}.\end{aligned} \tag{10.35}$$

The angular terms in the hydrogen Laplacian may then be rewritten as

$$\begin{aligned}\frac{1}{u^2}\frac{\partial^2}{\partial\phi^2} &= \frac{1}{u^2}\left(\frac{\partial^2}{\partial\beta^2} - 2\frac{\partial^2}{\partial\phi\partial\chi} - \frac{\partial^2}{\partial\chi^2}\right);\\ \frac{1}{v^2}\frac{\partial^2}{\partial\phi^2} &= \frac{1}{v^2}\left(\frac{\partial^2}{\partial\gamma^2} + 2\frac{\partial^2}{\partial\phi\partial\chi} - \frac{\partial^2}{\partial\chi^2}\right).\end{aligned} \tag{10.36}$$

Finally, we obtain

$$2r\,\Delta^2 = \sum_{i=1}^{4}\frac{\partial^2}{\partial u_i^2} + 2\frac{u^2 - v^2}{u^2 v^2}\frac{\partial^2}{\partial\phi\partial\chi} - \frac{u^2 + v^2}{u^2 v^2}\frac{\partial^2}{\partial\chi^2}. \tag{10.37}$$

This result can be replaced by the SU(4) Laplacian under the condition that the wave function is not dependent on the difference angle χ:

$$\frac{\partial}{\partial\chi}\,|\psi_n\rangle \equiv 0. \tag{10.38}$$

This condition puts a constraint on angular momenta in the u_1, u_2 and u_3, u_4 planes of the four-dimensional oscillator. Since χ is the difference between the polar angles in both planes, we have

$$\frac{\partial}{\partial\beta} = \frac{\partial}{\partial\gamma}, \tag{10.39}$$

or, in coordinate terms,

$$u_1 \frac{\partial}{\partial u_2} - u_2 \frac{\partial}{\partial u_1} = u_3 \frac{\partial}{\partial u_4} - u_4 \frac{\partial}{\partial u_3}. \tag{10.40}$$

This constraint breaks the SU(4) symmetry into two coupled SU(2) oscillators. In the next section, we will see how this gives rise to the SO(4) symmetry of the hydrogen atom. For now, let us continue with the Schrödinger equation for hydrogen. By multiplying both sides of the equation with $2r$ and applying the constraint, we finally obtain

$$\frac{-\hbar^2}{2m} \sum_{i=1}^{4} \frac{\partial^2}{\partial u_i^2} |\psi_n\rangle - \frac{Ze^2}{2\pi\varepsilon_0} |\psi_n\rangle = -\frac{Z^2 me^4}{8h^2{\varepsilon_0}^2 n^2} \left(u_1^2 + u_2^2 + u_3^2 + u_4^2\right) |\psi_n\rangle. \tag{10.41}$$

This is precisely the four-dimensional oscillator equation, with the following equivalences for frequencies and energy:

$$\begin{aligned} \frac{1}{2} m\omega^2 &\Longleftrightarrow \frac{Z^2 me^4}{8h^2{\varepsilon_0}^2 n^2}; \\ \hbar\omega \sum_{i=1}^{4} \left(n_i + \frac{1}{2}\right) &\Longleftrightarrow \frac{Ze^2}{2\pi\varepsilon_0}. \end{aligned} \tag{10.42}$$

These identities reduce to a very simple relationship between the respective quantum numbers:

$$\frac{1}{2} \sum_{i=1}^{4} \left(n_i + \frac{1}{2}\right) = n. \tag{10.43}$$

The remarkable feature of the KS transformation is the appearance of the radius as a common denominator to the kinetic and potential energy parts. By multiplying this denominator the singularity of the Coulomb hole in the origin, $r = 0$, is lifted and a much more tractable differential equation of oscillator type is obtained.

As noted earlier, the SU(4) symmetry is reduced because of the angular momentum constraint of Eq. (10.40). As a result, not all combinations of the n_i quantum numbers are allowed. In the next section, we will consider the symmetry reduction from SU(4) to SU(2) × SU(2), which ultimately will reproduce the hydrogen quantum numbers.

10.3 SHATTERING SU(4) SYMMETRY

The constraint introduced in Eq. (10.40) affects the SU(4) symmetry. Let $\{\hat{U}_x, \hat{U}_y, \hat{U}_z\}$ represent the three Lie generators of the SU(2) subgroup acting in the subspace formed by $\{u_1, u_2\}$, and similarly $\{\hat{V}_x, \hat{V}_y, \hat{V}_z\}$ represent the three Lie generators of the SU(2) subgroup acting in the subspace formed by $\{u_3, u_4\}$. The constraint requires the equality

$$\hat{U}_z = \hat{V}_z. \tag{10.44}$$

In view of this equality, generators from complementary subspaces might not commute. For example,

$$\left[\hat{U}_y, \hat{V}_z\right] = i\hat{U}_x. \tag{10.45}$$

This is not compatible with SU(4) symmetry, because this symmetry requires the two subspaces to be independent, and thus their generators should certainly commute. As a result, the constraint gives rise to an extra quantum condition that limits the quantum states of SU(4). To determine this condition, we first define creation and annihilation operators for boson states along the four degrees of freedom of SU(4). So $a_1^\dagger$ creates an excitation of the oscillator according to the coordinate u_1, and a_4 annihilates a boson excitation along u_4, and so on. The $\hat{U}_z$ and $\hat{V}_z$ generators may be expressed in these operators as indicated in Chapter 8:

$$\begin{aligned}\hat{U}_z &= \frac{\hbar}{2i}\left(a_1^\dagger a_2 - a_2^\dagger a_1\right);\\ \hat{V}_z &= \frac{\hbar}{2i}\left(a_3^\dagger a_4 - a_4^\dagger a_3\right).\end{aligned} \tag{10.46}$$

Since these operators rotate the boson excitations, it is convenient to change the basis to spinor form, which diagonalizes these operators. Following the procedure of Chapter 8, we define the corresponding spinors as follows:

$$\begin{aligned}a_\alpha^\dagger &= \frac{1}{\sqrt{2}}\left(a_1^\dagger + ia_2^\dagger\right);\\ a_\beta^\dagger &= \frac{1}{\sqrt{2}}\left(a_1^\dagger - ia_2^\dagger\right);\end{aligned} \tag{10.47}$$

$$\begin{aligned}b_\alpha^\dagger &= \frac{1}{\sqrt{2}}\left(a_3^\dagger + ia_4^\dagger\right);\\ b_\beta^\dagger &= \frac{1}{\sqrt{2}}\left(a_3^\dagger - ia_4^\dagger\right).\end{aligned} \tag{10.48}$$

We can do the same for the annihilation operators:

$$\begin{aligned}a_\alpha &= \frac{1}{\sqrt{2}}(a_1 - ia_2);\\ a_\beta &= \frac{1}{\sqrt{2}}(a_1 + ia_2);\end{aligned} \tag{10.49}$$

$$\begin{aligned}b_\alpha &= \frac{1}{\sqrt{2}}(a_3 - ia_4);\\ b_\beta &= \frac{1}{\sqrt{2}}(a_3 + ia_4).\end{aligned} \tag{10.50}$$

In this new basis, the generators become

$$\begin{aligned}\hat{U}_z &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\alpha - a_\beta^\dagger a_\beta\right);\\ \hat{V}_z &= \frac{\hbar}{2}\left(b_\alpha^\dagger b_\alpha - b_\beta^\dagger b_\beta\right).\end{aligned} \tag{10.51}$$

As explained in Appendix H, a general quantum state of SU(4) may be written as a boson excitation along all four degrees of freedom and is defined by the integer quantum numbers p, q, r, and s:

$$\left|p, q, r, s\right\rangle = \frac{1}{\sqrt{p!q!r!s!}}\left(a_\alpha^\dagger\right)^p \left(a_\beta^\dagger\right)^q \left(b_\alpha^\dagger\right)^r \left(b_\beta^\dagger\right)^s |0\rangle. \tag{10.52}$$

Applying the angular momentum operators of Eq. (10.51) to this state yields

$$\begin{aligned}\hat{U}_z\left|p, q, r, s\right\rangle &= \frac{\hbar}{2}(p-q)\left|p, q, r, s\right\rangle;\\ \hat{V}_z\left|p, q, r, s\right\rangle &= \frac{\hbar}{2}(r-s)\left|p, q, r, s\right\rangle.\end{aligned} \tag{10.53}$$

The equality in Eq. (10.44) thus yields an extra quantum condition:

$$p - q = r - s. \tag{10.54}$$

This equation may be reordered to yield

$$p + s = q + r. \tag{10.55}$$

This seemingly very simple rearrangement has a far-reaching group theoretical consequence. It defines another SU(2) × SU(2) cross-section, formed by the respective operator pairs $\{a_\alpha^\dagger, b_\beta^\dagger\}$ and $\{a_\beta^\dagger, b_\alpha^\dagger\}$. The basic generators for both subgroups are denoted as the respective vectors $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$, and they are given by

$$\begin{aligned} \hat{S}_x &= \frac{\hbar}{2}\left(a_\alpha^\dagger b_\beta + b_\beta^\dagger a_\alpha\right); \\ \hat{S}_y &= \frac{\hbar}{2i}\left(a_\alpha^\dagger b_\beta - b_\beta^\dagger a_\alpha\right); \\ \hat{S}_z &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\alpha - b_\beta^\dagger b_\beta\right); \end{aligned} \tag{10.56}$$

$$\begin{aligned} \hat{T}_x &= \frac{\hbar}{2}\left(a_\beta^\dagger b_\alpha + b_\alpha^\dagger a_\beta\right); \\ \hat{T}_y &= \frac{\hbar}{2i}\left(a_\beta^\dagger b_\alpha - b_\alpha^\dagger a_\beta\right); \\ \hat{T}_z &= \frac{\hbar}{2}\left(a_\beta^\dagger a_\beta - b_\alpha^\dagger b_\alpha\right). \end{aligned} \tag{10.57}$$

We can easily check that these operators form SU(2) algebras, and that the commutators of the $\hat{S}$ and $\hat{T}$ components vanish. The constraint in Eq. (10.55) does not introduce a mixing between the two groups; it only requires the number operators for both to be the same. In this way, this is a genuine symmetry breaking of SU(4) in a product of two separate SU(2) factors. Accordingly, the quantum states can be rewritten as

$$|p, q, r, s\rangle = |p, s\rangle\,|q, r\rangle. \tag{10.58}$$

As is explained in Appendix H, the SU(2) eigenfunctions can be characterized by a spin momentum quantum number j, which is given by

$$j = \frac{1}{2}\left(p + s\right) = \frac{1}{2}\left(q + r\right). \tag{10.59}$$

So, the symmetry restriction on the quantum states means that only SU(2) states with the same j can be transformed into the hydrogen wave functions. This is exactly what was obtained in Chapter 9, in which the SO(4) states were constructed as products of SU(2) states with the same j-value. Let us now introduce this condition into Eq. (10.43) for the n principal quantum number of hydrogen:

$$n = \frac{1}{2}\left(p + q + r + s + 2\right) = 2j + 1. \tag{10.60}$$

With this result, which reproduces Eq. (9.159) exactly, we are now prepared to leave the cozy world of compact invariance groups and head for the open-ended noninvariance groups of Part III.

PART THREE

SPECTRUM-GENERATING SYMMETRIES

11 The SO(2,1) group

Jacob left Beersheba, and went toward Haran. He came to the place and stayed there that night, because the sun had set. Taking one of the stones of the place, he put it under his head and lay down in that place to sleep. And he dreamed, and behold, there was a ladder set up on the earth, and the top of it reached to heaven; and behold, the angels of God were ascending and descending on it!
–Genesis 28:10–19.

The *radial wave equation* of the hydrogen atom contains a deep symmetry that relates eigenstates with different principal quantum number n, but equal orbital quantum number l. Under this peculiar symmetry, any nl-orbital can thus be turned into any other $n'l$-orbital (with $n \neq n'$ and l fixed, see Figure 11.1). Such a symmetry is truly *spectrum generating*, because its ladder operators run through the entire energy spectrum, from eigenstate to eigenstate.

Given that the energy E changes during such transformations, the Hamiltonian $\hat{\mathscr{H}}$ no longer remains invariant under these operations (i.e., $\hat{\mathscr{H}}$ does not commute with all the generators of the group), and the symmetry group is therefore called a *noninvariance group*, as will be explained in §11.1.

The spectrum-generating symmetry of the radial Schrödinger equation corresponds to the *noncompact* pseudo-orthogonal group SO(2,1). In §11.2, we will examine how this symmetry is related to its *compact* sister, the spherical rotation group SO(3). As a result of their similar structure, the algebraic treatment of the $\mathfrak{so}(2,1)$ algebra will closely resemble the approach used for the $\mathfrak{so}(3)$ algebra in Chapter 5.

FIGURE 11.1 Consider a deck of playing cards. There is exactly one card of any rank (1–10) in each suit (♡, ◇, ♣, or ♠). Let the suit represent the value of l (s, p, d, or f) and let the rank stand for the quantum number n (1–10). Then, with the help of the $\mathfrak{so}(2,1)$ generators, we can turn the cards of a given suit into one another, but cannot relate cards of different suits. The 2♠, 5♠, and 7♠ gardeners, depicted in this picture, are thus related to one another under the SO(2,1) group, but they can never rise in social class to become soldiers (♣), courtiers (◇), or members of the royal family (♡). In orbital language, any nl-orbital can be turned into any other $n'l$-orbital *with the same value of* l.

One of the fundamental differences between the SO(3) and SO(2,1) groups, to be outlined in §11.2, is that the unitary irreducible representations (or *unirreps*) of the $\mathfrak{so}(2,1)$ algebra are *infinite-dimensional*, whereas the unirreps of $\mathfrak{so}(3)$ are *finite-dimensional*. This sounds reasonable, given that for each value of l there are an infinite number of nl-orbitals (with $n = l + 1 \to \infty$) that form a basis for an infinite-dimensional representation of the $\mathfrak{so}(2,1)$ Lie algebra.

For this, we will have to investigate how the $\mathfrak{so}(2,1)$ algebra emerges in the case of the radial equation in §11.3. To this aim, a *realization* of the $\mathfrak{so}(2,1)$ generators $\hat{T}_1$, $\hat{T}_2$, and $\hat{T}_3$ will be given in terms of the radius $\hat{r}$ and conjugate momentum $\hat{p}_r$. By a clever *scaling transformation*, the radial Schrödinger equation can then be turned into an eigenvalue problem for one of the $\mathfrak{so}(2,1)$ generators, and the Bohr formula for the energy levels can be derived in an algebraic way, without recourse to the traditional differential equation approach in terms of series solutions.

Our examination of the SO(2,1) group will be based on the book chapter by Shi-Hai Dong on the *Lie algebra's SU(2) and SU(1,1)*.[1] This is also in line with the treatment by Adams and colleagues, and their notation will be used throughout.[2]

---

[1] S.-H. Dong. *Factorization Method in Quantum Mechanics.* Dordrecht: Springer, 2007, pp. 17–32.

[2] B. G. Adams, J. Čížek, and J. Paldus. "Representation Theory of so(4, 2) for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods." *International Journal of Quantum Chemistry* 21.1 (1982), pp. 153–171 and B. G. Adams, J. Čížek, and J. Paldus. "Lie Algebraic Methods and Their Application to Simple Quantum Systems." In: *Advances in Quantum Chemistry*, vol. 19. Ed. Per-Olov Löwdin. San Diego: Academic Press, 1988, pp. 1–84, reproduced in B. G. Adams, J. Čížek, and J. Paldus. "Lie Algebraic Methods and Their Application to Simple Quantum Systems." In: *Dynamical Groups and Spectrum-Generating Algebras*, vol. 1. Eds. Y. Ne'eman, A. Bohm, and A. O. Barut. Singapore: World Scientific Publishers, 1988, pp. 103–208. See also B. G. Adams. *Algebraic Approach to Simple Quantum Systems.* Berlin: Springer-Verlag, 1994, and

11.1 THE ROAD TO NONINVARIANCE GROUPS

Before we can deal with the SO(2,1) group, we have to acquaint ourselves with the notion of *noninvariance groups*. Although symmetry groups had been used in quantum mechanics since the groundbreaking work of Eugene Wigner and Weyl during the 1920s,[3] the idea to look for noninvariance groups under which *all* energy eigenstates of a system would fall in one irreducible representation had to wait another 40 years. As a matter of fact, the use of noninvariance groups was pioneered by Wigner as early as 1939 in relation to his study of the inhomogeneous Lorentz group.[4] Yet, it was only around 1965 that noninvariance groups first attracted the attention of a handful of high-energy physicists.

11.1.1 Historical prelude

Noninvariance groups made their first appearances in elementary particle physics. After Gell-Mann and Ne'eman's introduction of the SU(3) flavor group,[5] numerous theoretical physicists embarked on a quest for broader groups that would enable them to classify ever-larger numbers of elementary particles in so-called *supermultiplets*. Notorious examples were the SU(4) group, which took the newly discovered charm quark into account, or the spin-dependent SU(6) group, which accounted for both SU(3) flavor and SU(2) spin. As George Sudarshan noted with respect to the latter group, "its multiplet structure [was] remarkable."[6] No less than fifty-six baryons fell into one supermultiplet (or *56-plet*), whereas the mesons were seen to constitute a *35-plet*, both of which could be identified as irreducible representations of the spin-flavor SU(6) group.

But with the introduction of these symmetry groups, new problems inevitably surfaced. According to orthodox group theory, when different particles are grouped in one multiplet of a given symmetry group, they should be *degenerate* (read: have the same mass). From this point of view, the isospin SU(2) group was fairly accurate in nature because the proton and neutron have nearly equal mass. But, even within the SU(3) multiplets, significant differences in mass were observed. As a general rule, the higher the symmetry proposed, the more approximate it became. Sudarshan summarized the situation with the following words:

> While the multiplet structure is thus striking, there is a problem. If we thought of the group as an invariance group, we would have expected that the particles which

---

J. Čížek and J. Paldus. "An Algebraic Approach to Bound States of Simple One-Electron Systems." *International Journal of Quantum Chemistry* 12.5 (1977), pp. 875–896.

[3] E. P. Wigner. *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra*. New York: Academic Press, 1959, translated from the original E. P. Wigner. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*. Braunschweig: F. Vieweg & Sohn, 1931, and H. Weyl. *The Theory of Groups and Quantum Mechanics*. New York: Dover Publications, 1950, translated from the original H. Weyl. *Gruppentheorie und Quantenmechanik*. Leipzig: S. Hirzel, 1931.

[4] E. P. Wigner. "On Unitary Representations of the Inhomogeneous Lorentz Group." *Annals of Mathematics* 40.1 (1939), pp. 149–204.

[5] M. Gell-Mann and Y. Ne'eman. *The Eightfold Way*. Boulder, CO: Westview Press, 2000.

[6] E. C. G. Sudarshan. "What Are Elementary Particles Made Of?" In: *Symposia on Theoretical Physics and Mathematics*, vol. 6. Ed. Alladi Ramakrishnan. New York: Plenum Press, 1968, p. 61.

> constitute a single representation all have equal mass. On the other hand, if the group is not an invariance group, then it is not clear why the particles should constitute the representation of the group. There was a departure from the precise equality of masses for the SU(2) multiplets, but these differences were always blamed on the electromagnetic interaction. We thus have a paradoxical situation: On the one hand, the group seems to be basic in predicting the particle multiplets; on the other hand, the group does not seem to be an invariance group for the system.[7]

Instead of using these badly broken symmetries, the idea originated of looking for another kind of symmetry group—one that could relate all the particles in one multiplet, yet without forcing them to be degenerate. This is where *noninvariance groups* came into play. Scientists postulated a rather small symmetry group for the Hamiltonian and then embedded this group in a larger noninvariance group.[8]

Soon after the first successes in elementary particle physics, noninvariance groups began to be applied in the field of *atomic physics*. Sudarshan, Asim Orhan Barut, Yossef Dothan, Gell-Mann, and Ne'eman (among others), who had first pioneered the use of noninvariance groups in high-energy physics, now turned to a study of some familiar potentials, such as the *rigid rotator*, the *harmonic oscillator* (in three or more dimensions), and the *hydrogen atom*. In the latter case, the different (*ground* and *excited*) energy eigenstates were related to one another under the noninvariance group. Although these investigations were not of great *intrinsic* interest, both the harmonic oscillator and the hydrogen atom provided excellent test cases for noninvariance groups, and their study helped immensely in achieving a better understanding of their defining characteristics.

Noninvariance groups have been termed differently by various scientists: Sudarshan introduced the term *noninvariance group* (this term will also be adopted in the present chapter).[9] Gell-Mann, Ne'eman, and Dothan, on the other hand, preferred the name of *spectrum-generating groups*,[10] whereas Barut referred to them as *dynamical groups*, thus causing some confusion with our own use of that term for the SO(4) group (cf. Chapter 9, *vide infra*).[11]

Not only did terminological confusions abound during the 1960s and 1970s, the very concept of a *noninvariance group* was also seldom defined in a precise manner. As Octavio Castaños et al. noted, "[I]n contrast to the well-known prescription defining the elements of the symmetry algebra, … no unique prescription for the dynamical

---

[7] Sudarshan, "What are Elementary Particles Made of?," p. 61.

[8] J. G. Kuriyan and E. C. G. Sudarshan. "Noninvariance Groups in Particle Physics." *Physical Review* 162.5 (1967), pp. 1650–1662.

[9] N. Mukunda, L. O'Raifeartaigh, and E. C. G. Sudarshan. "Characteristic Noninvariance Groups of Dynamical Systems." *Physical Review Letters* 15.26 (1965), pp. 1041–1044. See also E. C. G. Sudarshan. "Concluding Remarks." In: *Non-Compact Groups in Particle Physics. Proceedings of the 1966 Conference held at the University of Wisconsin, Milwaukee*. Ed. Y. Chow. New York: Benjamin, 1966. pp. 207–212.

[10] Y. Dothan, M. Gell-Mann, and Y. Ne'eman. "Series of Hadron Energy Levels as Representations of Non-compact Groups." *Physics Letters* 17.2 (1965), pp. 148–151.

[11] A. O. Barut and C. Fronsdal. "On Non-compact Groups: II. Representations of the 2 + 1 Lorentz Group." *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* 287.1411 (1965), pp. 532–548.

[*sic*] group generators is [generally] provided."[12] Some definitions are therefore in place before we start our exploration of the SO(2,1) group and its relation to the hydrogen atom.

11.1.2 Invariance groups

We will distinguish between two kinds of groups: *invariance groups* and *noninvariance groups*. The invariance groups will be further classified as *geometrical groups* and *dynamical groups*.

Geometrical groups

Geometrical groups describe symmetries based on invariance properties of space and time. Recall that a Hamiltonian $\hat{\mathscr{H}}$ is said to be *invariant* under a Lie group $\mathcal{G}$ (or to have the *symmetry* $\mathcal{G}$) when it commutes with all the generators $\hat{X}_i$ of the corresponding Lie algebra $\mathfrak{g}$:

$$\left[\hat{\mathscr{H}}, \hat{X}_i\right] = 0. \tag{11.1}$$

The generators $\hat{X}_i$ are *Hermitian* operators ($\hat{X}_i^\dagger = \hat{X}_i$) that correspond to the different *constants of motion* X_i of the system under consideration.[13] The Lie algebra $\mathfrak{g}$, spanned by the generators $\hat{X}_i$, is called the *symmetry algebra*, or *degeneracy algebra* of the symmetry group $\mathcal{G}$, and is *closed* under commutation. That is to say, the commutator of any two generators $\hat{X}_i$ and $\hat{X}_j$ can be written as a linear combination of the generators of the Lie algebra:

$$\left[\hat{X}_i, \hat{X}_j\right] = \sum_{k=1} f_{ijk}\hat{X}_k \in \mathfrak{g}, \tag{11.2}$$

where f_{ijk} are the different *structure constants*. As a consequence of Eq. (11.1), *degeneracies* will arise in the energy eigenspectrum. That is, if $|\psi\rangle$ is an eigenstate of $\hat{\mathscr{H}}$ with energy E, then $\hat{X}_i|\psi\rangle$ is also an eigenstate of $\hat{\mathscr{H}}$ with the same energy E since $\hat{\mathscr{H}}\hat{X}_i|\psi\rangle = \hat{X}_i\hat{\mathscr{H}}|\psi\rangle = E\hat{X}_i|\psi\rangle$.

The *Casimir operators* of the Lie algebra $\mathfrak{g}$, denoted $\hat{C}_m$ in this chapter,[14] commute with all the generators $\hat{X}_i$:

$$\left[\hat{C}_m, \hat{X}_i\right] = 0. \tag{11.3}$$

Since the Hamiltonian operator commutes with all the generators (cf. Eq. (11.1)), it can be expressed in terms of the Casimir operators. That is, either $\hat{\mathscr{H}}$ turns out to be

---

[12] O. Castaños, A. Frank, and R. Lopez-Pena. "Noether's Theorem and Dynamical Groups in Quantum Mechanics." *Journal of Physics A: Mathematical and General* 23.22 (1990), p. 5141–5151.

[13] In classical Lagrangian mechanics, these constants of motion are often referred to as *conserved charges* or *currents*. The connection between conserved charges and continuous symmetries has been given before by *Noether's theorem* (cf. §6.3.1).

[14] The index m refers to the *order* of the Casimir operator in the different $\hat{X}_i$. For example, the Casimir operator of the $\mathfrak{so}(3)$ algebra is usually expressed as $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$, with $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ the generators of $\mathfrak{so}(3)$. Since $\hat{L}^2$ is written in terms of the *second* orders of the $\hat{L}_i$, we denote it as $\hat{C}_2$ and call it a Casimir operator of order 2.

one of the Casimir operators $\hat{C}_m$ of $\mathcal{G}$, or it can be written as a *polynomial function* in the Casimir operators $\hat{C}_m$ of various orders:

$$\hat{\mathscr{H}} = f\left(\hat{C}_m\right) = \sum_i \epsilon_i \hat{C}_i + \sum_{i,j} \epsilon_{ij} \hat{C}_i \hat{C}_j + \ldots. \tag{11.4}$$

Most symmetry groups are *compact* and result in *finite-dimensional multiplets* for each energy eigenvalue E. Examples include the SO(3) and SO(4) groups, considered before in Chapters 5 and 9. The degenerate eigenstates of a particular multiplet can then be related to one another via the ladder operators of the Cartan-Weyl basis. That is, the eigenstates form a *basis* for a *finite-dimensional unirrep* of the symmetry group $\mathcal{G}$.

Dynamical groups

A *dynamical symmetry* (not to be confounded with an *internal symmetry*) refers to any symmetry larger than the geometrical one. The Hamiltonian of the hydrogen atom, for instance, is invariant under the SO(4) group, which goes beyond the familiar rotational symmetry of other central potentials (Chapter 9).

Since dynamical groups are just another kind of invariance group, all the generators $\hat{X}_i$ of the corresponding Lie algebra $\mathfrak{g}$ commute with the Hamiltonian $\hat{\mathscr{H}}$ of the system (*vide supra*). Note, however, that for dynamical symmetries, the group generators typically commute with the *total* Hamiltonian, rather than with its (*kinetic* and *potential*) components separately, as is the case for most symmetry groups (cf. Chapter 9). Dynamical groups are usually *compact*, and the multiplets thus obtained remain *finite-dimensional* and *degenerate*.

11.1.3 Noninvariance groups

In the previous chapters, many invariance groups have been brought into play to explain the *energy degeneracy* of a system; that is, degenerate states can be transformed into one another under the action of the generators of the invariance group. It would be interesting to find a group capable of rationalizing not only the energy degeneracy for each level, but also the entire (discrete or continuous) *energy spectrum* (or *mass spectrum*) of the system—that is to say, a group with generators that link states with *different* energy eigenvalues, thus yielding the *energy level formula* as well as the energetic *ordering* of the levels.[15] To this aim, we will need the concept of a *noninvariance group*.

In this case, at least some of the generators $\hat{X}_i$ of the corresponding Lie algebra $\mathfrak{g}$ will no longer commute with the Hamiltonian. The commutator $\left[\hat{\mathscr{H}}, \hat{X}_i\right]$ is, however, an operator in the *universal enveloping algebra* $U(\mathfrak{g})$ of $\mathfrak{g}$, where $U(\mathfrak{g})$ is defined as the algebra of all polynomials in the generators $\hat{X}_i$[16]:

$$\left[\hat{\mathscr{H}}, \hat{X}_i\right] = f\left(\hat{X}_i\right) \in U(\mathfrak{g}). \tag{11.5}$$

---

[15] O. Sinanoğlu. "Remarks on Dynamical and Noncompact Groups in Physics and Chemistry." *International Journal of Quantum Chemistry* 7.S7 (1973), pp. 45–52.

[16] Every Casimir operator $\hat{C}_m$ of a Lie algebra $\mathfrak{g}$ is thus an element of the universal enveloping algebra $U(\mathfrak{g})$.

In some cases, $\hat{\mathscr{H}}$ can be written as one of the generators $\hat{X}_i \in \mathfrak{g}$. This is the case for the Hamiltonian of the radial Schrödinger equation described in this chapter. In other cases, $\hat{\mathscr{H}}$ is more complicated and belongs to $U(\mathfrak{g})$. Every noninvariance group (NIG) is required to contain as a subgroup the largest invariance group (IG) that can be built from the constants of the motion of the system under consideration[17]:

$$\text{NIG} \supset \text{IG}. \tag{11.6}$$

All the subgroup multiplets are therefore contained in one supermultiplet of the noninvariance group, which restricts the Casimir operators $\hat{C}_m$ of noninvariance group to constant *numbers.*[18]

Since the Hamiltonian no longer commutes with every generator $\hat{X}_i$, the energy is no longer conserved. This raises the possibility of altering the energy of the system by repeated application of raising and lowering operators. The entire *energy spectrum* of states can thus be generated, and it is for this reason that the corresponding Lie algebra $\mathfrak{g}$ is often referred to as a *spectrum-generating algebra* (SGA), with the symmetry algebra (SA) as a subalgebra:

$$\text{SGA} \supset \text{SA}. \tag{11.7}$$

Noninvariance groups can be both *compact* or *noncompact.* As a general rule, a finite (infinite) number of states calls for a compact (noncompact) group description. In the case of the hydrogen atom, for instance, noncompact groups are required to relate the infinite number of energy levels and states $|nlm\rangle$.

Noncompact groups give rise to *infinite-dimensional multiplets.* In the words of Dothan, "one poses a problem of embedding all the spaces of states which are irreducible under the symmetry algebra in a space which is irreducible under a larger Lie algebra, namely, the SGA."[19] All the energy eigenstates thus form a *basis* for a single *infinite-dimensional unirrep* of the noninvariance group.[20] Although invariance groups can be exploited to calculate the *matrix elements* for transitions connecting degenerate states, noninvariance groups can yield expressions for the various *transition probabilities* between states of different energy.

11.2 THE PSEUDO-ORTHOGONAL GROUP SO(2,1)

11.2.1 From SO(3) to SO(2,1)

The SO(3) group is the group of transformations that leaves a spherical surface invariant. Let a function $\mathscr{F}$ be defined as

$$\mathscr{F} = x^2 + y^2 + z^2 - a^2. \tag{11.8}$$

The equation of a sphere, with radius a and centered at the origin, is then given by

$$\mathscr{F} = 0. \tag{11.9}$$

---

[17] A. Simoni, F. Zaccaria, and B. Vitale. "Dynamical Symmetries as Function Groups on Dynamical Spaces." *Il Nuovo Cimento A* 51.2 (1967), pp. 448–460.

[18] Ibid., p. 450.

[19] Y. Dothan. "Finite-Dimensional Spectrum-Generating Algebras." *Physical Review D* 2.12 (1970), p. 2944.

[20] Noncompact groups can have *finite*-dimensional representations as well, but their *nonunitarity* renders them of little use in quantum mechanics. See Sinanoğlu, "Remarks on Dynamical and Noncompact Groups in Physics and Chemistry," p. 48.

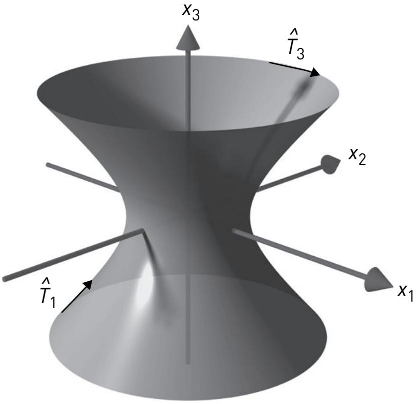


FIGURE 11.2 A three-dimensional *hyperboloid surface*. The x_1 and x_2 coordinates are *spacelike*, and the x_3 coordinate is *timelike*. The operator $\hat{T}_3$ induces infinitesimal rotations in the x_1x_2-plane, whereas $\hat{T}_1$ and $\hat{T}_2$ are *boost* operators that generate a flow in the x_2x_3- and x_1x_3-planes, respectively.

The angular momentum operators $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ induce infinitesimal rotations around the x-, y-, and z-axes, which leave the transformed point on the sphere. Their action on $\mathscr{F}$ is thus given by

$$\hat{L}_i\mathscr{F} = 0, \qquad \forall i = x, y, z. \tag{11.10}$$

Let us relabel the x-, y-, and z-coordinates as x_1, x_2, and x_3 to remove the link with the three Cartesian directions of coordinate space. Now, when one of the coordinates—say, x_3—is substituted by an imaginary coordinate ix_3, the equation of the sphere is turned into the equation of a *hyperboloid* (Figure 11.2)[21]:

$$\mathscr{F} = x_1^2 + x_2^2 - x_3^2 - a^2 = 0. \tag{11.11}$$

The cross-section of this surface for a given value of x_3, is a *circle* with radius $\sqrt{a^2 + x_3^2}$ in the x_1x_2-plane:

$$x_1^2 + x_2^2 = a^2 + x_3^2, \tag{11.12}$$

whereas in the planes x_1x_3 and x_2x_3, it is a *hyperbole*[22]:

$$x_1^2 - x_3^2 = a^2 - x_2^2; \tag{11.13}$$

$$x_2^2 - x_3^2 = a^2 - x_1^2. \tag{11.14}$$

We could thus construct the hyperboloid surface by rotating one of the hyperbolas in Eqs. (11.13) or (11.14) around the x_3-axis. In a sense, the hyperboloid surface can be viewed as a distorted sphere in which the north and south poles have been inflated to infinity.

Notice also that the hyperboloid expression in Eq. (11.11) is a two-dimensional analogue of the *Lorentz equation* of special relativity, in which the geodesic distance s between two events in four-dimensional Minkowski space-time is measured as

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2 = s^2. \tag{11.15}$$

---

[21] The general formula for this kind of three-dimensional surface is given by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

[22] The general formula for a hyperbole is given by

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Here, c is the velocity of light, which relates the units of space coordinates (in meters) and the time coordinate (in seconds), but reduces to unity on our hyperboloid surface so that x_3 and x_1, x_2 are measured in the same units. In this sense, we can say that the new x_3-coordinate in Eq. (11.11), which has opened up the sphere to infinity, is *timelike*, whereas x_1 and x_2 are *spacelike* coordinates.

11.2.2 The $\mathfrak{so}(2,1)$ Lie algebra

The operators that leave the hyperboloid invariant are directly obtained from the conventional angular momentum operators by applying the x_3/ix_3 substitution. As an example, the $\hat{L}_1$ operator becomes

$$\hat{L}_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \quad \rightarrow \quad \hat{L}_1 = -\hbar\left(x_2\frac{\partial}{\partial x_3} + x_3\frac{\partial}{\partial x_2}\right).$$

While in SO(3), $\hat{L}_x$ generates *rotations* on the sphere around the x-axis, the new form $\hat{L}_1$ generates a *flow* on the hyperboloid in the x_2x_3-plane (Figure 11.2). The operator $\hat{L}_1$ is still *anti-Hermitian* however (i.e., $\hat{L}_1^\dagger = -\hat{L}_1$). To procure Hermiticity, we divide the $\hat{L}_1$ operator through i, and similarly for the conversion of the $\hat{L}_2$ operator. The resulting canonical operators are denoted as $\hat{T}$-components and are then given by

$$\begin{aligned} \hat{T}_1 &= -\frac{\hbar}{i}\left(x_2\frac{\partial}{\partial x_3} + x_3\frac{\partial}{\partial x_2}\right); \\ \hat{T}_2 &= \frac{\hbar}{i}\left(x_3\frac{\partial}{\partial x_1} + x_1\frac{\partial}{\partial x_3}\right); \\ \hat{T}_3 &= \frac{\hbar}{i}\left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right). \end{aligned} \tag{11.16}$$

Clearly, the $\hat{T}_3$ operator still equals $\hat{L}_z$, because cross-sections perpendicular to x_3 remain circles. The other operators $\hat{T}_1$ and $\hat{T}_2$ are *boost operators* along the x_2x_3 and x_1x_3 hyperboles. The substitution x_3/ix_3 gives rise to a nontrivial sign change in the commutation relations. To compare:

$$\begin{aligned} \left[\hat{L}_x, \hat{L}_y\right] &= i\hbar\hat{L}_z; & \left[\hat{T}_1, \hat{T}_2\right] &= -i\hbar\hat{T}_3; \\ \left[\hat{L}_y, \hat{L}_z\right] &= i\hbar\hat{L}_x; & \left[\hat{T}_2, \hat{T}_3\right] &= i\hbar\hat{T}_1; \\ \left[\hat{L}_z, \hat{L}_x\right] &= i\hbar\hat{L}_y; & \left[\hat{T}_3, \hat{T}_1\right] &= i\hbar\hat{T}_2. \end{aligned} \tag{11.17}$$

The sign or *signature* of the commutation relations is different. For the $\mathfrak{so}(3)$ algebra it is $\{+,+,+\}$, whereas for the $\hat{T}$ algebra it is $\{-,+,+\}$. The Lie algebra formed by the $\hat{T}_i$ operators is therefore denoted $\mathfrak{so}(2,1)$, signifying it has two spacelike, or closed, components and one timelike, or open, component. The corresponding symmetry group SO(2,1), which in fact represents the plane Lorentz symmetry, is called a *noncompact* symmetry because one of the coordinates involved is unbounded. That is, although x_1 and x_2 are bounded, $x_3 \in]-\infty,+\infty[$ is obviously not.

11.2.3 The Cartan-Weyl basis of $\mathfrak{so}(2,1)$

For the $\mathfrak{so}(3)$ algebra, the Casimir operator is the square of the angular momentum vector $\hat{\mathbf{L}}$. In the noncompact $\mathfrak{so}(2,1)$ algebra, the equivalent of the angular momentum is the $\hat{\mathbf{T}}$ vector. In view of the sign difference of the commutators, the corresponding Casimir operator is given by[23]

$$\hat{T}^2 = \hat{T}_3^2 - \hat{T}_1^2 - \hat{T}_2^2. \tag{11.18}$$

This operator can be shown to commute with the three generators $\hat{T}_1$, $\hat{T}_2$, and $\hat{T}_3$ of the Lie algebra:

$$\left[\hat{T}^2, \hat{T}_i\right] = 0, \qquad \forall i = 1,2,3. \tag{11.19}$$

As an example, we can work out the commutators of each of the three terms of $\hat{T}^2$ with $\hat{T}_1$:

$$\begin{aligned}\left[\hat{T}_3^2, \hat{T}_1\right] &= \hat{T}_3\hat{T}_3\hat{T}_1 - \hat{T}_1\hat{T}_3\hat{T}_3 \\ &= \hat{T}_3\hat{T}_1\hat{T}_3 + i\hbar\hat{T}_3\hat{T}_2 - \hat{T}_3\hat{T}_1\hat{T}_3 + i\hbar\hat{T}_2\hat{T}_3 \\ &= i\hbar\left(\hat{T}_2\hat{T}_3 + \hat{T}_3\hat{T}_2\right);\end{aligned} \tag{11.20}$$

$$\left[\hat{T}_1^2, \hat{T}_1\right] = 0; \tag{11.21}$$

$$\begin{aligned}\left[\hat{T}_2^2, \hat{T}_1\right] &= \hat{T}_2\hat{T}_2\hat{T}_1 - \hat{T}_1\hat{T}_2\hat{T}_2 \\ &= \hat{T}_2\hat{T}_1\hat{T}_2 + i\hbar\hat{T}_2\hat{T}_3 - \hat{T}_2\hat{T}_1\hat{T}_2 + i\hbar\hat{T}_3\hat{T}_2 \\ &= i\hbar\left(\hat{T}_2\hat{T}_3 + \hat{T}_3\hat{T}_2\right).\end{aligned} \tag{11.22}$$

Combining the left-hand sides of these equations yields

$$\left[\hat{T}^2, \hat{T}_1\right] = 0. \tag{11.23}$$

An eigenstate can now be characterized by two scalar values, denoted $|Q,m\rangle$, where Q represents the value of the Casimir operator $\hat{T}^2$ and m refers to the value of the $\hat{T}_3$ component; hence,[24]

$$\begin{aligned}\hat{T}^2\,|Q,m\rangle &= Q\hbar^2\,|Q,m\rangle; \\ \hat{T}_3\,|Q,m\rangle &= m\hbar\,|Q,m\rangle.\end{aligned} \tag{11.24}$$

Following the same procedure as for SO(3), we now ask the question: What are the boundaries of these quantum labels? To this aim, ladder operators are defined in exactly the same way as before. The upward shift is represented by $\hat{T}_+$, and the downward one by $\hat{T}_-$, where

$$\hat{T}_\pm = \hat{T}_1 \pm i\hat{T}_2. \tag{11.25}$$

---

[23] The Casimir operator for $\mathfrak{so}(2,1)$ can be found as follows. Let $\hat{J}_1 = i\hat{T}_1$, $\hat{J}_2 = i\hat{T}_2$, and $\hat{J}_3 = \hat{T}_3$. We can easily verify that the $\hat{J}_i$ form an $\mathfrak{so}(3)$ algebra. Their Casimir operator is therefore $\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_3^2$. When written in terms of $\hat{T}$, we obtain the corresponding Casimir operator for the $\mathfrak{so}(2,1)$ algebra: $\hat{T}^2 = \hat{T}_3^2 - \hat{T}_1^2 - \hat{T}_2^2$. Notice that despite the notation, $\hat{T}^2$ does not denote the ordinary scalar product of a vector $\hat{\mathbf{T}}$ with itself. Instead, the scalar product is modified by the metric of the operator space.

[24] Notice that in view of the noncyclic symmetry of the commutation relations in Eq. (11.17), the choice of a $\hat{T}_i$ component is no longer completely equivalent. The sets of commuting operators $\{\hat{T}^2, \hat{T}_1\}$ or $\{\hat{T}^2, \hat{T}_2\}$ are more suitable for a study of the *continuum* (or *scattering*) states of the hydrogen atom, whereas the set $\{\hat{T}^2, \hat{T}_3\}$ will prove more useful in our study of the *bound* states of the hydrogen atom.

Notice that $\hat{T}_+$ and $\hat{T}_-$ are Hermitian adjoints of each other. They form the *Cartan-Weyl basis* $\{\hat{T}_3, \hat{T}_+, \hat{T}_-\}$ of $\mathfrak{so}(2,1)$, along with the Cartan generator $\hat{T}_3$. This gives rise to the following set of commutation relations:

$$\left[\hat{T}_+, \hat{T}_-\right] = -2\hbar \hat{T}_3; \tag{11.26}$$

$$\left[\hat{T}_3, \hat{T}_\pm\right] = \pm\hbar \hat{T}_\pm. \tag{11.27}$$

With the help of Eq. (11.27), it is easily seen that $\hat{T}_+$ and $\hat{T}_-$ indeed act as ladder operators in m:

$$\begin{aligned} \hat{T}_3 \hat{T}_\pm |Q,m\rangle &= \left(\left[\hat{T}_3, \hat{T}_\pm\right] + \hat{T}_\pm \hat{T}_3\right) |Q,m\rangle \\ &= \left(\pm\hbar \hat{T}_\pm + m\hbar \hat{T}_\pm\right) |Q,m\rangle \\ &= (m \pm 1)\, \hbar \hat{T}_\pm |Q,m\rangle. \end{aligned} \tag{11.28}$$

In Figure 11.3, these operators are presented in a Cartan-Weyl diagram based on the $\hat{T}_3$ operator. The Casimir operator can be easily expressed in the shift operators as follows:

$$\hat{T}^2 = \hat{T}_3^2 - \frac{1}{2}\left[\hat{T}_+ \hat{T}_- + \hat{T}_- \hat{T}_+\right]. \tag{11.29}$$

This expression can be further rewritten as

$$\hat{T}^2 = \hat{T}_3^2 - \hat{T}_+ \hat{T}_- - \hbar \hat{T}_3, \tag{11.30}$$

or also as

$$\hat{T}^2 = \hat{T}_3^2 - \hat{T}_- \hat{T}_+ + \hbar \hat{T}_3. \tag{11.31}$$

The expectation values of the ladder operator terms in these expressions can be shown to be real positive or zero:

$$\langle Q,m|\, \hat{T}_+ \hat{T}_- \,|Q,m\rangle = \left\langle \hat{T}_- Q,m \middle| \hat{T}_- Q,m \right\rangle \geq 0, \tag{11.32}$$

$$\langle Q,m|\, \hat{T}_- \hat{T}_+ \,|Q,m\rangle = \left\langle \hat{T}_+ Q,m \middle| \hat{T}_+ Q,m \right\rangle \geq 0, \tag{11.33}$$

where we have used the Hermitian conjugation between $\hat{T}_+$ and $\hat{T}_-$. From this follows

$$\begin{aligned} \langle Q,m| \left(\hat{T}^2 - \hat{T}_3^2 + \hbar \hat{T}_3\right) |Q,m\rangle &= \left(Q - m^2 + m\right) \hbar^2 \\ &= -\langle Q,m|\, \hat{T}_+ \hat{T}_- \,|Q,m\rangle \\ &\leq 0; \end{aligned} \tag{11.34}$$

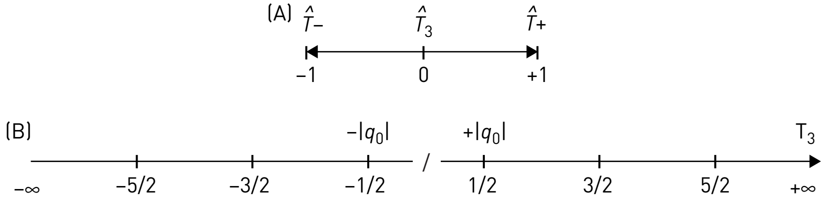


FIGURE 11.3 Root diagram of the $\mathfrak{so}(2,1)$ algebra. The $\hat{T}_3$ operator lies in the center of the graph and is used to *label* the eigenkets of a given multiplet. The $\hat{T}_+$ and $\hat{T}_-$ operators act as *shift* operators.

$$\begin{aligned}\langle Q,m|\left(\hat{T}^2-\hat{T}_3^2-\hbar\hat{T}_3\right)|Q,m\rangle &= \left(Q-m^2-m\right)\hbar^2\\ &= -\langle Q,m|\,\hat{T}_-\hat{T}_+\,|Q,m\rangle \qquad (11.35)\\ &\leq 0.\end{aligned}$$

In line with the earlier treatment of SO(3) we now redefine the Q-parameter by the parameter expression $j\left(j+1\right)$:

$$Q\equiv j\left(j+1\right). \qquad (11.36)$$

Note that for a given value of Q, two solutions for j are possible: either j or $-j-1$. Inserting this parameter definition in the previous inequalities yields two requirements that j and m have to fulfill:

$$Q-m^2+m=j\left(j+1\right)-m^2+m=\left(j+m\right)\left(j-m+1\right)\leq 0; \qquad (11.37)$$

$$Q-m^2-m=j\left(j+1\right)-m^2-m=\left(j-m\right)\left(j+m+1\right)\leq 0. \qquad (11.38)$$

Now let us consider a representation space based on the symmetry quantities j and m. The inequality requirements define regions of existence in this space. These regions are delineated by four straight lines, corresponding to the zeros of the four monomials in Eqs. (11.37) and (11.38). In Figure 11.4, these lines are indicated by lowercase letters. We get

$$\begin{aligned}&\text{line a}: j-m=0;\\ &\text{line b}: j+m=0;\\ &\text{line c}: j-m+1=0; \qquad (11.39)\\ &\text{line d}: j+m+1=0.\end{aligned}$$

Above line a, the monomial $j-m$ is positive, whereas below the line it is negative. This line thus divides the representation space into two parts with opposite signs for the

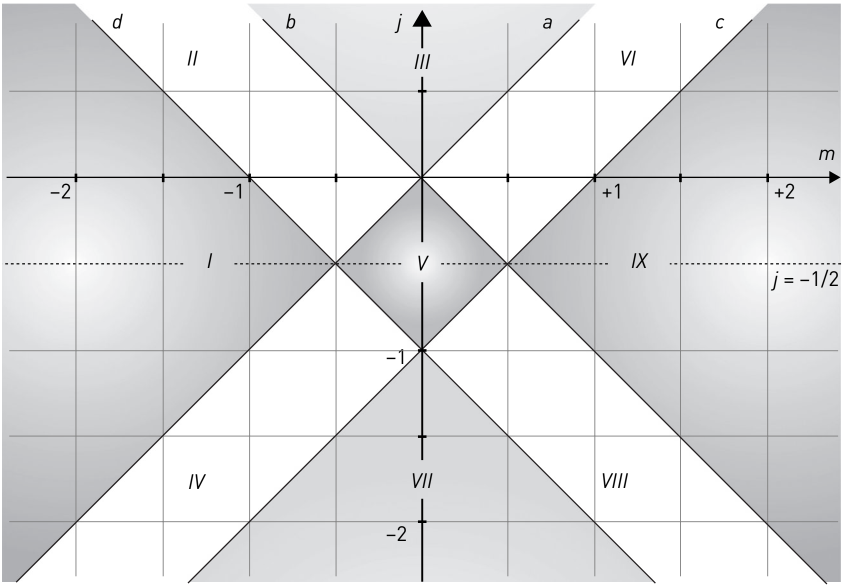


FIGURE 11.4 The sign requirements of Eqs. (11.37) and (11.38) divide the *jm*-plane in nine sectors.

determining conditions. The set of four lines together forms a mesh that partitions the plane in no less than nine sectors, denoted by Roman numerals. An overview of the signs of the four monomials in these sectors is provided in Table 11.1.

From the table, it is clear that there are only three sectors in which both requirements of Eqs. (11.37) and (11.38) are fulfilled: I, V, and IX. Hence, these are the only regions with allowed combinations of j and m. They are marked as SO(2,1) regions in the table. In addition, there are two regions in which both conditions are greater than or equal to zero—namely, III and VII. These are marked as SO(3). We will return to these later.

The dashed line in Figure 11.4 at $j=-1/2$ is a line of reflection symmetry. A given point $(j,\ m)$ is mirrored through this symmetry line in a point $(-j-1,\ m)$. Because the values j and $-j\ -\ 1$ yield the same Q-value, both points refer to the same eigenket—meaning, only one half of the representation space has to be considered. For the case of SO(2,1), we consider the lower part, with $j \leq -1/2$. A detailed view of this part is shown in Figure 11.5. As can be seen from the diagram, this space

Table 11.1 Sign analysis of the monomial delimiters of the sectors in the SO(2,1) weight diagram.

	$j+m$	$j-m+1$	$\times$	$j-m$	$j+m+1$	$\times$	
I	−	+	−	+	−	−	SO(2,1)
II	−	+	−	+	+	+	
III	+	+	+	+	+	+	SO(3)
IV	−	+	−	−	−	+	
V	−	+	−	−	+	−	SO(2,1)
VI	+	+	+	−	+	−	
VII	−	−	+	−	−	+	SO(3)
VIII	−	−	+	−	+	−	
IX	+	−	−	−	+	−	SO(2,1)

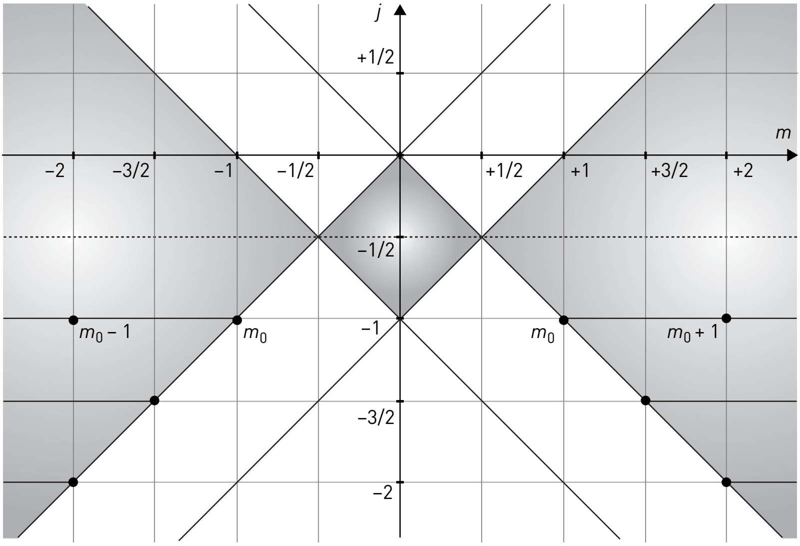


FIGURE 11.5 Detailed view of the SO(2,1) sectors in the *jm*-plane.

contains two principal branches: a positive one corresponding to sector IX and a negative branch in the opposite sector, I. These branches enclose the central diamond sector V. Now consider a point $(j,\ m)$ in the sector corresponding to the positive branch. Applying the $\hat{T}_-$ operator to this point displaces it on a horizontal line to the left over a unit distance, reaching $(j,\ m-1)$. Similarly the $\hat{T}_+$ operator displaces it to the right to $(j,\ m+1)$. Although the displacements to the right can go on forever, the displacements to the left are bounded by the b line, corresponding to $j+m=0$. Let $(j,\ m_0)$ be the point with the smallest m-value. From Eq. (11.30), we can infer

$$\begin{aligned}\hat{T}_+\hat{T}_-\,|Q,m_0\rangle &= \left(-\hat{T}^2+\hat{T}_3^2-\hbar\hat{T}_3\right)|Q,m_0\rangle \\ &= (j+m_0)\,(-j+m_0-1)\,\hbar^2\,|Q,m_0\rangle.\end{aligned} \tag{11.40}$$

Since the lowering operator must destroy the m_0 eigenket, $\hat{T}_-\,|Q,m_0\rangle=0$, it is required that $m_0=-j$. Hence, the representations corresponding to the positive branch start precisely at the b line and run horizontally in unit steps, corresponding to

$$m=-j,-j+1,\ -j+2,\ldots. \tag{11.41}$$

Similarly for the negative branch, the representation starts on the a line and runs to the left in unit steps to $-\infty$; hence,

$$m=j,j-1,j-2,j-3,\ldots. \tag{11.42}$$

Both branches are disconnected because the gap between the two cannot be bridged by the ladder operators. Note that until now, it was not required that j itself be integer or half-integer. In the subsequent quantum problem, however, only representations with integer values of j will appear.

A special case occurs for $(j,\ m)$ combinations that fall inside the central diamond region. In this case, ladder operators extend these points both to the left and to the right, and generate truly unbound representations.

11.2.4 SO(3) revisited

As already indicated, the $(j,\ m)$ representation space also accommodates the SO(3) representations. In this case, the defining conditions are very similar to the SO(2,1) conditions in Eqs. (11.37) and (11.38), except that the expressions should now be greater or equal to zero:

$$(j+m)\,(j-m+1)\geq 0; \tag{11.43}$$

$$(j-m)\,(j+m+1)\geq 0. \tag{11.44}$$

As can be seen from Table 11.1, there are two sectors for which both requirements are fulfilled: III and VII. The mirror symmetry of the representation space permutes these sectors; hence, they refer to identical representations. In line with the previous treatment of SO(3), we will consider sector III, with $j\geq 0$ (Figure 11.6). Now $(j,\ m)$ points in this sector are displaced by the ladder operators horizontally in between the boundary lines a and b. Hence, in contrast to the SO(2,1) representations, the SO(3) manifolds have a lower and an upper boundary. Because the ladder operators cannot displace components outside the boundaries, the monomials $(j+m)$ and

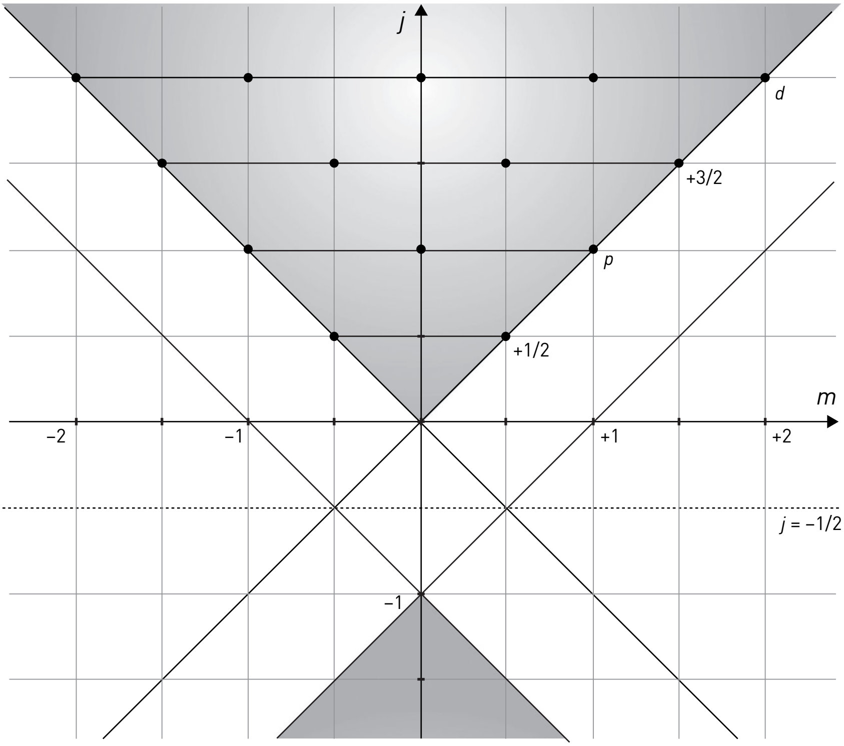


FIGURE 11.6 Detailed view of the SO(3) sectors in the *jm*-plane.

$(j-m)$ must vanish on the left and right boundaries, respectively. This implies that the components with minimal and maximal m-values must be situated on the boundary lines:

$$m_{min} = -j; \tag{11.45}$$

$$m_{max} = +j. \tag{11.46}$$

Or,

$$m_{max} - m_{min} = 2j. \tag{11.47}$$

Since the distance between the two boundaries is bridged by an integer number of unit steps, it follows that $2j$ must be integer, precisely as we derived in Chapter 5.

11.3 HYDROGENIC REALIZATION OF THE SO(2,1) GROUP

Around the first half of the twentieth century, group theory seemed to be "the wave of the future."[25] One of the first scientists to champion its importance was British astronomer Sir Arthur Eddington (1882–1944) who (as early as 1934) likened the theory of groups with "a super-mathematics in which the operations are as unknown as the quantities they operate on."[26] Eddington meant by this that *abstract* group theory deals with a set of *abstract* elements that combine according to an

---

[25] F. J. Dyson. "Applications of Group Theory in Particle Physics." *SIAM Review* 8.1 (1966), p. 1.

[26] A. Eddington. *New Pathways in Science (Messenger Lectures delivered at Cornell in April and May, 1934)*. Cambridge: Cambridge University Press, 1935, pp. 255–277.

FIGURE 11.7 "Beware the Jabberwock!"[27]

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgabe.

"Jabberwocky" is considered one of the greatest nonsense poems written by Lewis Carroll. As the late Sir Arthur Eddington observed, "it is certainly descriptive of some kind of activity; but what the actors are, and what kind of actions they are performing, remain an inscrutable mystery." In a similar vein, Eddington likened group theory to a *Jabberwocky* of "unknowable actors executing unknowable actions," and awaiting a physical realization in terms of concrete operators.

abstract multiplication rule, the (physical) nature of which is unknown and irrelevant mathematically (cf. Chapter 2, §2.2).[28]

The power of group theory, then, lays in the fact that one abstract group could be *realized* in an infinite number of ways by different sets of concrete elements that make their appearance in a range of physical applications (Figure 11.7). The SO(4) group, for example, was seen to occur in both the classical Kepler and quantum mechanical Coulomb problem. The abstract SO(2,1) symmetry, on the other hand, makes its appearance in the *radial Schrödinger equation* of the hydrogen atom. In this section, a concrete set of operators will be associated with each abstract basis vector $\hat{T}_i$ of the $\mathfrak{so}(2,1)$ Lie algebra to obtain a *hydrogenic realization* of the $\mathfrak{so}(2,1)$ algebra.

11.3.1 The radial Schrödinger equation

We start by expressing the kinetic energy of a single particle with mass m in terms of its spherical components as (cf. Eq. (5.146))

$$\hat{T}_{kin} = \frac{1}{2m}\,\hat{p}_r^2 + \frac{1}{2mr^2}\,\hat{L}^2. \tag{11.48}$$

---

[27] Lewis Carroll. *Jabberwocky and Other, Nonsense, Collected Poems.* Ed. Gillian Beer. London: Penguin. The poem was included in the 1871 edition of: *Through the Looking Glass and What Alice Found There.*

[28] "We must therefore seek a knowledge which is neither of actors nor of actions," continued Eddington, "but of which the actors and actions are a vehicle. The knowledge we can acquire is knowledge of a *structure* or pattern contained in the actions" [emphasis added]. Ibid., p. 256.

Here, $\hat{p}_r$ is the *radial momentum operator*, which is given by

$$\hat{p}_r = \frac{1}{r}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} - \frac{i\hbar}{r} = \frac{\hbar}{i}\left(\frac{\partial}{\partial r} + \frac{1}{r}\right). \tag{11.49}$$

The square of this operator reads

$$\hat{p}_r^2 = -\hbar^2\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right). \tag{11.50}$$

The total Hamiltonian for the hydrogen atom is obtained by adding to Eq. (11.48) the potential energy operator, yielding

$$\hat{\mathscr{H}} = \hat{T}_{kin} - \frac{e^2}{4\pi\epsilon_0 r}. \tag{11.51}$$

Since the potential energy operator only depends on the radius, the angular part of the kinetic energy operator can be solved separately. To this aim, the hydrogenic wave functions ψ are written in spherical coordinates and separated as a product of a radial function $R_{n,l}$ and a spherical harmonic Y_l^m:

$$\psi\,(r,\theta,\phi) = R_{n,l}\,(r)\,Y_l^m\,(\theta,\phi). \tag{11.52}$$

As we saw in Chapter 5, the angular part gives rise to SO(3) angular momentum states characterized by the orbital quantum number l, with

$$\hat{L}^2\,|l,m\rangle = l\,(l+1)\,\hbar^2\,|l,m\rangle. \tag{11.53}$$

Substitution of this result in the operator part allows us to factorize out the angular part completely, yielding the so-called *radial equation* of the hydrogen atom:

$$\hat{\mathscr{H}}_{rad} = \frac{1}{2m}\,\hat{p}_r^2 + \frac{l\,(l+1)\,\hbar^2}{2mr^2} - \frac{e^2}{4\pi\epsilon_0 r}. \tag{11.54}$$

The term $l\,(l+1)$ represents the “centrifugal” potential that provides an outward pressure on the electron cloud as a result of its rotatory motion. Only for $l=0$ (i.e., for s-type orbitals that present a static cloud) does this term vanish.

11.3.2 The spectrum-generating $\mathfrak{so}$ (2,1) algebra

The previous equation incorporates the $\mathfrak{so}(2,1)$ algebra as a *spectrum-generating* algebra, as will be shown here. The basic radial operators are the radius r itself and the conjugate momentum $\hat{p}_r$. With these, the following three operators are constructed:

$$\begin{aligned}
\hat{Q}_1 &= \frac{1}{2}\left(\frac{r\hat{p}_r^2}{a} - ar + \frac{b}{r}\right);\\
\hat{Q}_2 &= r\hat{p}_r;\\
\hat{Q}_3 &= \frac{1}{2}\left(\frac{r\hat{p}_r^2}{a} + ar + \frac{b}{r}\right).
\end{aligned} \tag{11.55}$$

Here, a and b are real scaling parameters that introduce extra degrees of freedom. Quite remarkably, these operators embody precisely the Lie algebraic commutators of the SO(2,1) group:

$$\begin{aligned}
\left[\hat{Q}_1,\hat{Q}_2\right] &= -i\hbar\hat{Q}_3;\\
\left[\hat{Q}_2,\hat{Q}_3\right] &= i\hbar\hat{Q}_1;\\
\left[\hat{Q}_3,\hat{Q}_1\right] &= i\hbar\hat{Q}_2.
\end{aligned} \tag{11.56}$$

The proof of this result again boils down to working out the commutators between the radial operators. These can be determined from Eqs. (11.49) and (11.50):

$$\begin{aligned}\left[r\hat{p}_r, r\hat{p}_r^2\right] &= i\hbar\, r\hat{p}_r^2 \\ \left[r\hat{p}_r, r\right] &= -i\hbar r \\ \left[r\hat{p}_r, \frac{1}{r}\right] &= i\hbar\frac{1}{r} \\ \left[r\hat{p}_r^2, r\right] &= -2i\hbar r\hat{p}_r.\end{aligned} \tag{11.57}$$

A few manipulations are now sufficient to express the radial Eq. (11.54) in terms of these basic $\hat{Q}$ operators. Let us first write the radial Schrödinger equation as $\left(\hat{\mathscr{H}} - E\right)|\Psi\rangle = 0$, where E is the total energy and $|\Psi\rangle$ represents a radial eigenstate $R_{n,l}(r)$:

$$\left(\frac{1}{2m}\hat{p}_r^2 + \frac{l(l+1)\,\hbar^2}{2mr^2} - \frac{e^2}{4\pi\epsilon_0 r} - E\right)|\Psi\rangle = 0. \tag{11.58}$$

Now we multiply (on the left) with mr:

$$\left(\frac{r\hat{p}_r^2}{2} + \frac{l(l+1)\,\hbar^2}{2r} - \frac{me^2}{4\pi\epsilon_0} - mrE\right)|\Psi\rangle = 0. \tag{11.59}$$

Subsequently, we introduce a real parameter a that is defined by

$$a^2 = -2mE. \tag{11.60}$$

We then divide the radial equation by a and obtain

$$\left[\frac{1}{2}\left(\frac{r\hat{p}_r^2}{a} + \frac{l(l+1)\,\hbar^2}{ar} + ar\right) - \frac{me^2}{4\pi\epsilon_0 a}\right]|\Psi\rangle = 0. \tag{11.61}$$

The operator part in this expression is now recast in a form that is immediately recognized as the $\hat{Q}_3$ operator, with

$$\frac{l(l+1)\,\hbar^2}{a} = b. \tag{11.62}$$

The parameter a can be incorporated as a scaling parameter into a rescaled radius R. The conjugated momentum, $\hat{P}_R$, is rescaled accordingly by $1/a$:

$$R = a\,r, \quad \hat{P}_R = \frac{1}{a}\hat{p}_r. \tag{11.63}$$

This yields

$$\left[\frac{1}{2}\left(R\hat{P}_R^2+\frac{l(l+1)\,\hbar^2}{R}+R\right)-\frac{me^2}{4\pi\epsilon_0 a}\right]|\Psi\rangle=0. \tag{11.64}$$

We should keep in mind that the parameter a is a function of the energy. Eq. (11.60) can be rewritten as

$$a=\pm\sqrt{-2mE}. \tag{11.65}$$

Since a is real, it is required that the total energy E be negative. So, we retrieve the quantization condition for bound hydrogen states. We further assume that a be positive, since it serves as a radial scaling parameter. The radial equation can then be rewritten in terms of the SO(2,1) operator as

$$\left(\hat{Q}_3-\frac{me^2}{4\pi\epsilon_0 a}\right)|\Psi\rangle=0. \tag{11.66}$$

The eigenstates of the radial equation are thus eigenstates of the $\hat{Q}_3$ operator:

$$\hat{Q}_3\,|\Psi\rangle=\frac{me^2}{4\pi\epsilon_0 a}\,|\Psi\rangle. \tag{11.67}$$

The total momentum operator can be expressed in radial operators as follows:

$$\begin{aligned}\hat{\mathbf{Q}}^2&=\hat{Q}_3^2-\hat{Q}_1^2-\hat{Q}_2^2\\&=\left(\hat{Q}_3-\hat{Q}_1\right)\left(\hat{Q}_3+\hat{Q}_1\right)-\left[\hat{Q}_3,\hat{Q}_1\right]-\hat{Q}_2^2\\&=ar\left(\frac{r\hat{p}_r^2}{a}+\frac{b}{r}\right)-i\hbar\hat{Q}_2-\hat{Q}_2^2\\&=R^2\hat{P}_R^2+ab-i\hbar R\hat{P}_R-R\hat{P}_R R\hat{P}_R\\&=ab,\end{aligned} \tag{11.68}$$

where in the final line we have made use of the commutator $[\hat{P}_R,R]=-i\hbar$, which leads to

$$R\hat{P}_R R\hat{P}_R=R^2\hat{P}_R^2-i\hbar R\hat{P}_R. \tag{11.69}$$

In this way, all variables vanish. The remaining constant term for the case of the hydrogen equation reduces to

$$ab=l(l+1)\,\hbar^2. \tag{11.70}$$

We thus harvest the important conclusion that the value associated with the squared total momentum $\hat{\mathbf{Q}}^2$ is equal to the value of the squared angular momentum $\hat{\mathbf{L}}^2$. This implies that the states of an infinite-dimensional SO(2,1) multiplet, characterized by a constant value of j, share the same value of l. The transformations, induced by the ladder operators $\hat{Q}_\pm=\hat{Q}_1\pm i\hat{Q}_2$, are therefore restricted to transformations that link states with the same value of l. An s-orbital, for example, can be transformed into any other s-orbital, but cannot be related to p-, d-, or f-orbitals.

Now we have to examine the two different SO(2,1) representations. As indicated, we chose the scaling factor a in Eq. (11.65) to be positive, so that R retains the characteristics of a distance. As a result, the eigenvalues for the $\hat{Q}_3$ operator in

Eq. (11.67) are found to be positive as well. This puts us in the positive branch of the SO(2,1) representations, corresponding to sector IX in Figure 11.5. We then have

$$\begin{aligned} j(j+1) &= l(l+1); \\ j &= -l-1; \\ m_0 &= -j = l+1. \end{aligned} \tag{11.71}$$

Hence, the Q_3 components start at $l+1$ and increase in steps of unity along the horizontal line in the diagram. The $\hat{Q}_+$ ladder operator takes the ground level $|j, m_0\rangle$ to the higher floors $|j, m_j\rangle$, with

$$m_j = m_0 + n_r = 1 + l + n_r \text{ with } n_r = 0,\ 1,\ 2,\ \ldots. \tag{11.72}$$

Here, n_r is the radial quantum number. It counts the number of radial nodes. Since the radial solutions were found to be eigenstates of $\hat{Q}_3$, we can also identify the eigenvalues with the m_j quantum numbers. This, finally, yields the quantized energy levels of the hydrogen atom:

$$\hbar m_j = \frac{me^2}{4\pi\epsilon_0 a} = \hbar(1+l+n_r). \tag{11.73}$$

Combining these expressions we obtain

$$\frac{me^2}{4\pi\epsilon_0\sqrt{-2mE}} = \hbar(1+l+n_r), \tag{11.74}$$

or

$$E = -\frac{me^4}{8h^2\epsilon_0^2 n^2} \quad \text{with} \quad n = 1+l+n_r = 1,\ 2,\ 3,\ \ldots. \tag{11.75}$$

If we had started by the (nonphysical) assumption that the radial scaling parameter would be negative, the m_j values would have been negative as well, and we would have situated the ladder in the negative branch, corresponding to sector I. This would not have altered the resulting expressions.

11.4 DYNAMICAL TREATMENT OF THE RADIAL WAVE EQUATION

Consider the positive branch, sector IX, in the Cartan-Weyl diagram (Figure 11.4). The Lie algebra shows that the $\hat{Q}_-$ operator destroys the lowest component $|j, m_0\rangle$ of an SO(2,1) manifold. If this is applied to the radial equation, the annihilation of the lowest component becomes

$$\left(\hat{Q}_1 - i\hat{Q}_2\right)|j, m_0\rangle = \left[\frac{1}{2}\left(R\hat{P}_R^2 - R + \frac{l(l+1)\hbar^2}{R}\right) - iR\hat{P}_R\right]|j, m_0\rangle = 0. \tag{11.76}$$

For simplicity let us restrict ourselves to the case $l = 0$, which also implies $j = -1$. The equation then reduces to

$$\begin{aligned} \frac{R}{2}\left[\hat{P}_R^2 - 1 - 2i\hat{P}_R\right]|j, m_0\rangle &= 0 \\ \left(\hat{P}_R - i\right)^2 |j, m_0\rangle &= 0. \end{aligned} \tag{11.77}$$

Introducing in this equation the explicit form of the radial momentum operator, we obtain a second-order differential equation in the radial wave function

$$\left(-\hbar^2\frac{\partial^2}{\partial R^2} - 2\hbar^2\frac{1}{R}\frac{\partial}{\partial R} - 1 - 2\hbar\frac{\partial}{\partial R} - 2\hbar\frac{1}{R}\right)|j, m_0\rangle = 0. \tag{11.78}$$

To solve this equation, we use the Ansatz:

$$\left|j, m_0\right\rangle = f \exp\left(-\frac{R}{\hbar}\right), \tag{11.79}$$

where f is a function of R. When this form is inserted into the differential equation, a simplified equation in f is obtained:

$$-\hbar^2\left[\frac{2}{R}\frac{\partial f}{\partial R} + \frac{\partial^2 f}{\partial R^2}\right]\exp\left(-\frac{R}{\hbar}\right) = 0, \tag{11.80}$$

which reduces to

$$\frac{2}{R}\frac{\partial f}{\partial R} = -\frac{\partial^2 f}{\partial R^2}. \tag{11.81}$$

By integrating this equation, we see that the function f can be either constant or equal to $1/R$. In the first case, the fundamental radial function of the $1s$ ground level reads

$$\Psi_{1s} = \exp\left(-\frac{R}{\hbar}\right), \tag{11.82}$$

which still has to be normalized. Recall that R is a scaled radius ar. For the lowest level $(n = 1)$, the value of a is given by

$$a = \frac{me^2}{2h\epsilon_0}. \tag{11.83}$$

Note that the a for the $1s$ ground level is related to the so-called *Bohr radius*, a_0, which is the atomic unit of length:

$$a_0 = \frac{h^2\epsilon_0}{\pi me^2} = \frac{\hbar}{a}. \tag{11.84}$$

We thus recover the familiar form of the $1s$ wave function as

$$\Psi_{1s} = \exp\left(-\frac{r}{a_0}\right). \tag{11.85}$$

In addition, we also learn that there may be a second solution given by $1/r \exp(-r/a_0)$. This solution brings us to the foundations of quantum mechanics. Although it is sound mathematically, it has to be rejected on physical grounds. As stated by Dirac, when the wave function approaches the origin as $1/r$, it implies that the probability of finding the particle at the origin is infinitely favored over all other positions, which is physically unacceptable.[29] Dirac went on to prove that the product $r\Psi$ should tend toward zero as one approaches the origin.

11.5 DIRAC'S HARMONIC OSCILLATOR REVISITED

Our first encounter with ladder operators that do not conserve energy occurred during the quantum mechanical treatment of the harmonic oscillator in Chapter 7. The symmetry group that hosts the creation and annihilation operators $a^\dagger$ and a clearly must be a noninvariance group, but what is the corresponding Lie algebra? So far, we have identified only one commutation relation. In the following, we put the energy parameter $\hbar\omega$ equal to unity. Hence,

$$\left[a, a^\dagger\right] = \mathbb{I}. \tag{11.86}$$

---

[29] P. A. M. Dirac. *The Principles of Quantum Mechanics.* Oxford: Clarendon Press, 1930, p. 156.

We thus have a structure with two generators and a constant $\mathbb{I}$, which may be identified as the Casimir operator of the entire space of operator states. The Hamiltonian, on the other hand, is neither of these. It brings into play a relation of a different nature—the symmetrized combination of creation and annihilation operators, which is known as the *anticommutator*. This is the opposite of a commutator and it is usually denoted by curly brackets:

$$\hat{\mathscr{H}} = \frac{1}{2}\left(a^\dagger a + aa^\dagger\right) = \frac{1}{2}\left\{a^\dagger, a\right\}. \tag{11.87}$$

We thus stumble across a combination of commutator and anticommutator brackets. Evidently, two more anticommutators come into play as well: $\{a, a\}$ and $\left\{a^\dagger, a^\dagger\right\}$. The combination of all these forms a structure that extends the Lie algebra to a *superalgebra*. The elements of this superalgebra are the creation and annihilation operators, and their anticommutators. The proper definitions are as follows:

$$\begin{aligned}
\hat{Q} &= \frac{1}{\sqrt{2}}a;\\
\hat{Q}^\dagger &= \frac{1}{\sqrt{2}}a^\dagger;\\
\hat{K}_3 &= \frac{1}{4}\left\{a^\dagger, a\right\};\\
\hat{K}_+ &= \frac{1}{4}\left\{a^\dagger, a^\dagger\right\};\\
\hat{K}_- &= \frac{1}{4}\{a, a\}.
\end{aligned} \tag{11.88}$$

Brackets in this superalgebra now appear in two varieties. Brackets in which both entries are odd powers of the Q operators are anticommutators whereas other brackets are commutators:

$$\begin{aligned}
\left\{\hat{Q}^\dagger, \hat{Q}\right\} &= 2\hat{K}_3 = \hat{\mathscr{H}};\\
\left\{\hat{Q}^\dagger, \hat{Q}^\dagger\right\} &= 2\hat{K}_+;\\
\left\{\hat{Q}, \hat{Q}\right\} &= 2\hat{K}_-;\\
\left[\hat{K}_3, \hat{Q}^\dagger\right] &= +\frac{1}{2}\hat{Q}^\dagger;\\
\left[\hat{K}_3, \hat{Q}\right] &= -\frac{1}{2}\hat{Q};\\
\left[\hat{K}_+, \hat{Q}^\dagger\right] &= 0;\\
\left[\hat{K}_+, \hat{Q}\right] &= -\hat{Q}^\dagger;\\
\left[\hat{K}_-, \hat{Q}^\dagger\right] &= +\hat{Q};\\
\left[\hat{K}_-, \hat{Q}\right] &= 0;\\
\left[\hat{K}_3, \hat{K}_\pm\right] &= \pm\hat{K}_\pm;\\
\left[\hat{K}_+, \hat{K}_-\right] &= -2\hat{K}_3.
\end{aligned} \tag{11.89}$$

This algebra is known as the *orthosymplectic* Lie superalgebra, $\mathfrak{osp}(1|2)$. Comparison with the SO(2,1) Lie algebra shows that the commutation relations for the $\hat{K}$ operators are identical to the commutation relationships for the $\hat{T}$-operators in Eq. (11.27). The $\mathfrak{so}(2,1)$ Lie algebra is thus a subgroup of the orthosymplectic superalgebra:

$$\mathfrak{so}(2,1) \subset \mathfrak{osp}(1|2). \tag{11.90}$$

The Casimir operator for this subgroup is defined as in Eq. (11.29):

$$\hat{K}^2 = \hat{K}_3^2 - \frac{1}{2}\left[\hat{K}_+\hat{K}_- + \hat{K}_-\hat{K}_+\right]. \tag{11.91}$$

Since the $\hat{K}_\pm$ ladder operators involve double excitations, the harmonic oscillator spectrum forms two separate multiplets of this SO(2,1) group: one multiplet consisting of all even excitations and one consisting of all odd excitations. As indicated in Appendix H, a normalized eigenstate of n excitations is given by

$$|\Psi_n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n\,|0\rangle. \tag{11.92}$$

Application of the $\hat{K}$ operators (in units of $\hbar\omega$) on this state yields

$$\begin{aligned} \hat{K}_3^2\,|\Psi_n\rangle &= \frac{\left(n+\frac{1}{2}\right)^2}{4}\,|\Psi_n\rangle; \\ \hat{K}_+\hat{K}_-\,|\Psi_n\rangle &= \frac{n\,(n-1)}{4}\,|\Psi_n\rangle; \\ \hat{K}_-\hat{K}_+\,|\Psi_n\rangle &= \frac{(n+2)\,(n+1)}{4}\,|\Psi_n\rangle. \end{aligned} \tag{11.93}$$

The resulting Casimir operator then yields

$$\hat{K}^2\,|\Psi_n\rangle = -\frac{3}{16}\,|\Psi_n\rangle. \tag{11.94}$$

Expressing this eigenvalue as $k_0\,(k_0-1)$ results in a value of $k_0 = 1/4$ or $k_0 = 3/4$. As opposed to the SO(2,1) subgroup, the full orthosymplectic supergroup allows single excitations as well and thus unites the even and odd multiplets into one eigenspace. The extended Casimir operator for this supermultiplet can be expressed as

$$\hat{C} = \hat{K}^2 + \left[\hat{Q},\hat{Q}^\dagger\right]. \tag{11.95}$$

Since the Q-commutator yields a constant (see Eq. (11.86)), this combined operator is merely a trivial extension of the SO(2,1) Casimir operator.

Interestingly, as Eq. (11.89) shows, the $\hat{Q}$ operators themselves transform irreducibly in the SO(2,1) subgroup. It can indeed be shown that these operators commute with the $\hat{K}^2$ Casimir operator. We have

$$\begin{aligned} \left[\hat{K}_3^2,\hat{Q}\right] &= \left[\hat{K}_3,\hat{Q}\right]\hat{K}_3 + \hat{K}_3\left[\hat{K}_3,\hat{Q}\right] \\ &= -\frac{1}{2}\hat{Q}\hat{K}_3 - \frac{1}{2}\hat{K}_3\hat{Q} = \frac{1}{4}\hat{Q} - \hat{Q}\hat{K}_3; \\ \left[\hat{K}_+\hat{K}_-,\hat{Q}\right] &= \left[\hat{K}_+,\hat{Q}\right]\hat{K}_- + \hat{K}_+\left[\hat{K}_-,\hat{Q}\right] \\ &= -\hat{Q}^\dagger\hat{K}_- \\ \left[\hat{K}_-\hat{K}_+,\hat{Q}]\right] &= \left[\hat{K}_-,\hat{Q}\right]\hat{K}_+ + \hat{K}_-\left[\hat{K}_+,\hat{Q}\right] \\ &= -\hat{K}_-\hat{Q}^\dagger = -\hat{Q} - \hat{Q}^\dagger\hat{K}_-. \end{aligned} \tag{11.96}$$

Combining these expressions finally yields the total commutator with the SO(2,1) Casimir operator:

$$\left[\hat{K}^2,\hat{Q}\right]=\frac{3}{4}\hat{Q}-\hat{Q}\hat{K}_3+\hat{Q}^\dagger\hat{K}_-=0, \tag{11.97}$$

and similarly for $\hat{Q}^\dagger$. Clearly the creation and annihilation operators form a spin 1/2 doublet in SO(2,1). As Tristan Hübsch pointed out, "amusingly thus, although the operators a and $a^\dagger$ are bosonic, the superalgebra relations use anticommutators for odd powers of a and $a^\dagger$ precisely as if they were fermionic, anticommuting annihilation and creation operators."[30]

---

[30] T. Hubsch. "Spectrum-Generating Superalgebra for Linear Harmonic Oscillators." In: *arXiv*. No. 1203.5103 [math-ph] (2012), p. 4.

12 The SO(4,2) group

The universe {which others call the Library} is composed of an indefinite and perhaps infinite number of hexagonal galleries, with vast air shafts between, surrounded by very low railings. From any of the hexagons one can see, interminably, the upper and lower floors. The distribution of the galleries is invariable. Twenty shelves, five long shelves per side, cover all the sides except two; their height, which is the distance from floor to ceiling, scarcely exceeds that of a normal bookcase. One of the free sides leads to a narrow hallway which opens onto another gallery, identical to the first and to all the rest.
–Jorge Luis Borges.[1]

In 1896, two students from New Zealand were awarded the 1851 Exhibition Scholarship to undertake university studies abroad. One was John Angus Erskine (1873–1960) and the other Ernest Rutherford (1871–1937). Rutherford went to Manchester, where he would discover the atomic nucleus. Erskine became an engineer and later left to the University of Canterbury, Christchurch, New Zealand, a munificent bequest to enable teaching staff to travel overseas and bring specialists in the same fields to lecture at the university.

In 1971, the names of the two scholarship holders appeared again in connection with our subject. In that year, theoretical physicist Asim Orhan Barut (1926–1994) from the University of Colorado was the visiting Erskine Fellow, and he also attended

---

[1] J. L. Borges. "The Library of Babel." In: *Labyrinths, Selected Stories and Other Writings by Jorge Luis Borges.* Eds. Donald A. Yates and James E. Irby. New York: New Directions, 1962, p. 51. The story was originally published in Borges' 1941 collection of stories *El Jardín de Senderos Que se Bifurcan.* Buenos Aires: Sur, 1941.

the Rutherford centennial symposium on the structure of matter, which was held in Christchurch from July 7 to 9, 1971. His visit to New Zealand yielded two important documents about the group structure of the Periodic Table. One is a contribution in the proceedings of the *Rutherford Symposium.*[2] The other is a little-known booklet containing notes of the lectures Barut gave in the same year as Erskine Fellow.[3]

In these lectures on "Dynamical Groups and Generalized Symmetries in Quantum Theory," Barut explained the construction of the noncompact Lie group SO(4,2), which is the covering spectrum-generating group for the hydrogen problem. The spectrum of bound states of hydrogen turns out to be one single representation of the group SO(4,2). This representation was later named the *baruton*,[4] because it can be viewed as a single particle that manifests itself in the Periodic Table of the elements. We will come back to this important point in the final chapter.

The SO(4,2) group unites all the knowledge about the hydrogen spectrum we have gathered so far in the preceding chapters. In Chapter 5, we introduced the spherical symmetry group SO(3), which describes the spatial symmetry of the angular equation and is applicable to any central field problem in three dimensions. In Chapter 9, we obtained the hyperspherical group SO(4), which turned out to explain the degeneracy of the hydrogen levels. Finally, in Chapter 11, the covering group SO(2,1) was obtained, which describes the dynamical symmetry of the radial Schrödinger equation for the hydrogen atom.

The combination of all these symmetries provides shift operators that allow us to run through the entire set of bound states of hydrogen; any state $|nlm\rangle$ can thus be obtained from the ground state $|100\rangle$ by the subsequent application of the appropriate ladder operators. The SO(3) group contains ladder operators that change the magnetic quantum number m; the SO(4) group allows us to shift the orbital quantum number l in a given n-shell; and last, but not least, the spectrum-generating group SO(2,1) yields shift operators that go up and down the principal quantum number n, thereby changing the energy.

The respective algebras taken together yield nine infinitesimal operators: six from the $\mathfrak{so}(4)$ Lie algebra ($\hat{L}_1$, $\hat{L}_2$, $\hat{L}_3$, $\hat{A}_1$, $\hat{A}_2$ and $\hat{A}_3$), and three more from the $\mathfrak{so}(2,1)$ algebra ($\hat{Q}_1$, $\hat{Q}_2$, and $\hat{Q}_3$). In order to *merge* both algebras to obtain a larger unified Lie algebra, it should be verified whether these nine operators form a *closed set* under commutation. As will be shown in this chapter, the nine operators do *not* close, and six additional operators have to be invoked to close the commutation relations. This completes the merging process and yields a six-dimensional $\mathfrak{so}(4,2)$ Lie algebra of order $r = 15$ and rank $l = 3$.

In this chapter we will examine how all these operators can be combined to yield the spectrum-generating symmetry group SO(4,2). In the first section, we will look

---

[2] A. O. Barut. "Group Structure of the Periodic System." In: *The Structure of Matter: Rutherford Centennial Symposium.* Ed. B. G. Wybourne. Christchurch, New Zealand: University of Canterbury Press, 1972, pp. 126–136.

[3] A. O. Barut. *Dynamical Groups and Generalized Symmetries in Quantum Theory* (*With Applications in Atomic and Particle Physics*) Christchurch, New Zealand: Bascands, 1972.

[4] C. E. Wulfman. "Dynamical Groups in Atomic and Molecular Physics." In: *Recent Advances in Group Theory and Their Application to Spectroscopy.* Ed. J. C. Donini. New York: Plenum Press, 1978, pp. 329–403.

at the construction of this symmetry group, along the lines set out by Barut. In the second section, we will present weight and root diagrams for the baruton of hydrogen states and consider the Casimir invariants. The final section is devoted to the moves of the ladder operators in a hydrogenic basis.

12.1 THE PSEUDO-ORTHOGONAL GROUP SO(4,2)

12.1.1 From SO(6) to SO(4,2)

The special orthogonal group in six dimensions, SO(6), corresponds to the group of rotations in six-dimensional Euclidean space $\mathbb{R}^6$, or—what is equivalent—the set of real 6 × 6 orthogonal matrices $\mathbb{R}$, which leave the quadratic form

$$\mathscr{F}(\mathbf{r}) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 = \mathbf{r}^{\mathrm{T}}\mathbf{r} \tag{12.1}$$

invariant, where $\mathbf{r}$ is the column vector $[x_1, x_2, x_3, x_4, x_5, x_6]^{\mathrm{T}}$ and $\mathbf{r}^{\mathrm{T}}$ is its dual transpose row vector. To see this, consider a general matrix transformation $\mathbb{R}\mathbf{r}$ that preserves the quadratic form $\mathscr{F}(\mathbf{r})$:

$$\mathscr{F}(\mathbb{R}\mathbf{r}) = (\mathbb{R}\mathbf{r})^{\mathrm{T}}(\mathbb{R}\mathbf{r}) = \mathbf{r}^{\mathrm{T}}\mathbb{R}^{\mathrm{T}}\mathbb{R}\mathbf{r} = \mathbf{r}^{\mathrm{T}}\mathbf{r} = \mathscr{F}(\mathbf{r}). \tag{12.2}$$

This holds only if $\mathbb{R}$ is orthogonal:

$$\mathbb{R}^{\mathrm{T}}\mathbb{R} = \mathbb{I}, \tag{12.3}$$

where $\mathbb{I}$ is the 6 × 6 identity matrix.

Definition 12.1 (The SO(6) Lie group): The set of proper rotation matrices

$$\mathrm{SO}(6) = \left\{ \text{real } 6 \times 6 \text{ matrices } \mathbb{R} : \begin{array}{l} \mathbb{R}^{\mathrm{T}}\mathbb{R} = \mathbb{R}\mathbb{R}^{\mathrm{T}} = \mathbb{I} \\ \det \mathbb{R} = 1 \end{array} \right\} \tag{12.4}$$

forms a group under matrix multiplication. This group is called the *special orthogonal group* in six dimensions and is denoted by the symbol SO(6). ■

The structure of the corresponding $\mathfrak{so}(6)$ algebra just builds on a generalization of the familiar angular momentum operators, which we already encountered in our examination of the $\mathfrak{so}(2)$, $\mathfrak{so}(3)$, and even $\mathfrak{so}(4)$ Lie algebras. The momenta may be conveniently labeled by a row and column index that refer to the entries in a 6 × 6 matrix. The operators $\hat{L}_{ab}$ are then given by the formula

$$\hat{L}_{ab} = \frac{\hbar}{i}\left(x_a \frac{\partial}{\partial x_b} - x_b \frac{\partial}{\partial x_a} \right), \qquad \forall a, b = 1 \to 6. \tag{12.5}$$

These operators are Hermitian and they form a basis for the $\mathfrak{so}(6)$ Lie algebra. The number of independent generators is easily found. Out of the thirty-six possible combinations of the indices a and b, six vanish in view of the fact that $\hat{L}_{aa} = 0$ (for $a = 1 \to 6$), which decreases the number of generators to thirty. Moreover, an interchange of the row and column indices corresponds to a sign change; hence,

$$\hat{L}_{ab} = -\hat{L}_{ba}. \tag{12.6}$$

As a result, only fifteen independent generators remain, the number of which is also given by the formula $n(n-1)/2$, where n is the dimension of the group (see Chapter 5):

$$\mathbb{L}=\begin{pmatrix} 0 & \hat{L}_{12} & \hat{L}_{13} & \hat{L}_{14} & \hat{L}_{15} & \hat{L}_{16} \\ & 0 & \hat{L}_{23} & \hat{L}_{24} & \hat{L}_{25} & \hat{L}_{26} \\ & & 0 & \hat{L}_{34} & \hat{L}_{35} & \hat{L}_{36} \\ & & & 0 & \hat{L}_{45} & \hat{L}_{46} \\ & & & & 0 & \hat{L}_{56} \\ & & & & & 0 \end{pmatrix}. \tag{12.7}$$

Notice that this matrix can be extended to the lower half using the antisymmetry relation in Eq. (12.6). Each of the $\hat{L}_{ab}$ operators generates a rotation in the $x_a x_b$-plane. The commutation relations are

$$\left[\hat{L}_{ab},\hat{L}_{ac}\right]=-\hbar^2\left(x_c\frac{\partial}{\partial x_b}-x_b\frac{\partial}{\partial x_c}\right)=i\hbar\hat{L}_{bc}. \tag{12.8}$$

The latter equation can be rewritten in a more formal way as

$$\left[\hat{L}_{ab},\hat{L}_{ac}\right]=g_{aa}i\hbar\hat{L}_{bc}, \qquad \forall a\neq b\neq c, \text{ no sum on } a. \tag{12.9}$$

The g factors are structure factors that determine the *metric* of the space. For the compact SO(6) group, the metric or curvature is positive everywhere; hence, the row of g's may be represented as $[+,+,+,+,+,+]$. That is, the *metric matrix* $\mathbb{G}=\operatorname{diag}(1,1,1,1,1,1)$:

$$\mathbb{G}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{12.10}$$

which can be identified with the 6×6 identity matrix $\mathbb{I}$.

12.1.2 Hydrogenic realization of the $\mathfrak{so}(4,2)$ Lie algebra

In Chapter 9, we identified the degeneracy group of the hydrogen atom as the SO(4) group. In the full spectrum-generating symmetry group, this group must be present and we can immediately take over the identification we made in Chapter 9:

$$\begin{aligned} \hat{L}_1=\hat{L}_{23}; \quad \hat{L}_2=\hat{L}_{31}; \quad \hat{L}_3=\hat{L}_{12}; \\ \hat{A}_1=\hat{L}_{14}; \quad \hat{A}_2=\hat{L}_{24}; \quad \hat{A}_3=\hat{L}_{34}. \end{aligned} \tag{12.11}$$

In analogy with Eq. (12.7), this correspondence can also be represented as:

$$\mathbb{L}\Longleftrightarrow\begin{pmatrix} 0 & \hat{L}_3 & -\hat{L}_2 & \hat{A}_1 \\ & 0 & \hat{L}_1 & \hat{A}_2 \\ & & 0 & \hat{A}_3 \\ & & & 0 \end{pmatrix}. \tag{12.12}$$

Note the negative sign for $\hat{L}_2$, which is identified by cyclic permutation as $\hat{L}_{31}$, and thus corresponds to $-\hat{L}_{13}$. Their defining commutation relations are repeated here:

$$\left[\hat{L}_i,\hat{L}_j\right]=i\varepsilon_{ijk}\hat{L}_k; \tag{12.13}$$

$$\left[\hat{L}_i, \hat{A}_j\right] = i\varepsilon_{ijk}\hat{A}_k; \tag{12.14}$$

$$\left[\hat{A}_i, \hat{A}_j\right] = i\varepsilon_{ijk}\hat{L}_k. \tag{12.15}$$

In addition, we have to include into the covering group the dynamical $\mathfrak{so}(2,1)$ algebra of the radial wave equation introduced in Chapter 11, which contains three more operators $\hat{Q}_i$ $(i = 1 \to 3)$. Recall that the hydrogenic wave functions can be written as a product of a radial function with a spherical harmonic function:

$$\psi(r,\theta,\phi) = R_{n,l}(r)\, Y_l^m(\theta,\phi). \tag{12.16}$$

Since the $\hat{Q}_i$ operators act on the radial part of the hydrogenic wave functions only, whereas the angular momentum components operate on the angular part, both operators are expected to commute:

$$\left[\hat{L}_i, \hat{Q}_i\right] = 0. \tag{12.17}$$

Therefore, we must accommodate the subalgebra of $\hat{Q}_i$ operators in a three-by-three block that has no common indices with the angular momentum block. To this end, we must extend the matrix with two additional indices, 5 and 6, to form a 456 block separate from the 123 block. The entries may be identified as

$$\hat{Q}_1 = \hat{L}_{46}; \quad \hat{Q}_2 = \hat{L}_{45}; \quad \hat{Q}_3 = \hat{L}_{56}. \tag{12.18}$$

When written in the form of Eq. (12.7), we obtain

$$\mathbb{L} \Longleftrightarrow \begin{pmatrix} 0 & \hat{L}_3 & -\hat{L}_2 & \hat{A}_1 & \cdot & \cdot \\ & 0 & \hat{L}_1 & \hat{A}_2 & \cdot & \cdot \\ & & 0 & \hat{A}_3 & \cdot & \cdot \\ & & & 0 & \hat{Q}_2 & \hat{Q}_1 \\ & & & & 0 & \hat{Q}_3 \\ & & & & & 0 \end{pmatrix}. \tag{12.19}$$

However since SO(2,1) is noncompact, the metric factors g_{55} and g_{66} have to be negative:

$$\begin{aligned} \left[\hat{Q}_1, \hat{Q}_2\right] &= -i\hbar\hat{Q}_3 & \to \quad \left[\hat{L}_{46}, \hat{L}_{45}\right] &= i\hbar g_{44}\hat{L}_{65} = -i\hbar\hat{L}_{56}; \\ \left[\hat{Q}_2, \hat{Q}_3\right] &= i\hbar\hat{Q}_1 & \to \quad \left[\hat{L}_{45}, \hat{L}_{56}\right] &= -i\hbar g_{55}\hat{L}_{46} = i\hbar\hat{L}_{46}; \\ \left[\hat{Q}_3, \hat{Q}_1\right] &= i\hbar\hat{Q}_2 & \to \quad \left[\hat{L}_{56}, \hat{L}_{46}\right] &= i\hbar g_{66}\hat{L}_{54} = i\hbar\hat{L}_{45}. \end{aligned} \tag{12.20}$$

We thus obtain a noncompact group, with metric $[+,+,+,+,-,-]$, that we have to denote as SO(4,2) instead of SO(6). The metric matrix is then written as $\mathbb{G} = \mathrm{diag}(1,1,1,1,-1,-1)$:

$$\mathbb{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \tag{12.21}$$

With this metric, the commutation relations in Eq. (12.9) give rise to the commutation Table 12.1. The $\hat{Q}$ operators contain two parameters, a and b, that we identified in the previous chapter as follows:

$$a = \sqrt{-2mE}; \tag{12.22}$$

$$b = \frac{l(l+1)\,\hbar^2}{a}. \tag{12.23}$$

This parametrization should raise some concern, because the operators of the algebra rely on the energy of the states on which they are acting. However, this problem can be solved in a consistent way by a uniform rescaling of the radial distance, as we shall now see. Since the radial and angular momenta in the expression for the $\hat{Q}$ operators are scaled by the same factor, we can reintroduce the total momentum $\hat{p}$ as follows:

$$\frac{r\hat{p}_r^2}{a} + \frac{b}{r} = \frac{1}{a}\left(r\hat{p}_r^2 + \frac{l(l+1)\,\hbar^2}{r}\right) = \frac{r\hat{p}^2}{a}. \tag{12.24}$$

In this way, the $\hat{Q}$ operators become

$$\begin{aligned} \hat{Q}_1 &= \frac{1}{2}\left(\frac{r\hat{p}^2}{a} - ar\right); \\ \hat{Q}_2 &= r\hat{p}_r; \\ \hat{Q}_3 &= \frac{1}{2}\left(\frac{r\hat{p}^2}{a} + ar\right). \end{aligned} \tag{12.25}$$

Note that the parameter b is now fully absorbed in the momentum operator. The parameter a can also be reset to unity by simply scaling the radius ar to R. This scaling will also replace the expressions $a^{-1}\hat{p}_r$ and $a^{-2}\hat{p}^2$ by $\hat{P}_R$ and $\hat{P}^2$, respectively. Hence, under this scaling, the $\hat{Q}$s are simplified to

$$\begin{aligned} \hat{Q}_1 &= \frac{1}{2}\left(R\hat{P}^2 - R\right); \\ \hat{Q}_2 &= R\hat{P}_R; \\ \hat{Q}_3 &= \frac{1}{2}\left(R\hat{P}^2 + R\right). \end{aligned} \tag{12.26}$$

Recall that the LRL vector contains the energy as a parameter:

$$\hat{\mathbf{A}} = \frac{1}{\sqrt{-2mE}}\left(\frac{1}{2}\hat{\mathbf{r}}\hat{p}^2 - \hat{\mathbf{p}}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right) + mE\hat{\mathbf{r}}\right). \tag{12.27}$$

Using the expression for a, this becomes

$$\hat{\mathbf{A}} = \frac{\hat{\mathbf{r}}\hat{p}^2}{2a} - \frac{\hat{\mathbf{p}}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)}{a} - \frac{1}{2}a\hat{\mathbf{r}}. \tag{12.28}$$

Here, too, a rescaling by a absorbs this parameter completely in the radial distance, and the LRL becomes

$$\hat{\mathbf{A}} = \frac{1}{2}\hat{\mathbf{R}}\hat{P}^2 - \hat{\mathbf{P}}\left(\hat{\mathbf{R}}\cdot\hat{\mathbf{P}}\right) - \frac{1}{2}\hat{\mathbf{R}}. \tag{12.29}$$

On the other hand, the angular momentum operators $\hat{L}_1$, $\hat{L}_2$, $\hat{L}_3$ remain invariant under this scaling. In this way, all parameter dependence can be eliminated from the operator algebra and we obtain nine momentum operators that are dependent only on the coordinates and associated momenta.

However, with six indices, a total of fifteen $\hat{L}_{ij}$ operators can be formed. The remaining six operators can be easily obtained by forming the required commutators

Table 12.1 Commutation table for the generators of the $\mathfrak{so}(4,2)$ Lie algebra in terms of generalized angular momentum operators $\hat{L}_{ab}$. The commutator $[\hat{X}_i, \hat{X}_j]$ is listed in the ith row and jth column.

	$\hat{L}_{23}$	$\hat{L}_{31}$	$\hat{L}_{12}$	$\hat{L}_{14}$	$\hat{L}_{24}$	$\hat{L}_{34}$	$\hat{L}_{15}$	$\hat{L}_{25}$	$\hat{L}_{35}$	$\hat{L}_{16}$	$\hat{L}_{26}$	$\hat{L}_{36}$	$\hat{L}_{46}$	$\hat{L}_{45}$	$\hat{L}_{56}$
$\hat{L}_{23}$	0	$i\hbar\hat{L}_{12}$	$-i\hbar\hat{L}_{31}$	0	$i\hbar\hat{L}_{34}$	$-i\hbar\hat{L}_{24}$	0	$i\hbar\hat{L}_{35}$	$-i\hbar\hat{L}_{25}$	0	$i\hbar\hat{L}_{36}$	$-i\hbar\hat{L}_{26}$	0	0	0
$\hat{L}_{31}$	$-i\hbar\hat{L}_{12}$	0	$i\hbar\hat{L}_{23}$	$-i\hbar\hat{L}_{34}$	0	$i\hbar\hat{L}_{14}$	$-i\hbar\hat{L}_{35}$	0	$i\hbar\hat{L}_{15}$	$-i\hbar\hat{L}_{36}$	0	$i\hbar\hat{L}_{16}$	0	0	0
$\hat{L}_{12}$	$i\hbar\hat{L}_{31}$	$-i\hbar\hat{L}_{23}$	0	$i\hbar\hat{L}_{24}$	$-i\hbar\hat{L}_{14}$	0	$i\hbar\hat{L}_{25}$	$-i\hbar\hat{L}_{15}$	0	$i\hbar\hat{L}_{26}$	$-i\hbar\hat{L}_{16}$	0	0	0	0
$\hat{L}_{14}$	0	$i\hbar\hat{L}_{34}$	$-i\hbar\hat{L}_{24}$	0	$i\hbar\hat{L}_{12}$	$-i\hbar\hat{L}_{31}$	$i\hbar\hat{L}_{45}$	0	0	$i\hbar\hat{L}_{46}$	0	0	$-i\hbar\hat{L}_{16}$	$-i\hbar\hat{L}_{15}$	0
$\hat{L}_{24}$	$-i\hbar\hat{L}_{34}$	0	$i\hbar\hat{L}_{14}$	$-i\hbar\hat{L}_{12}$	0	$i\hbar\hat{L}_{23}$	0	$i\hbar\hat{L}_{45}$	0	0	$i\hbar\hat{L}_{46}$	0	$-i\hbar\hat{L}_{26}$	$-i\hbar\hat{L}_{25}$	0
$\hat{L}_{34}$	$i\hbar\hat{L}_{24}$	$-i\hbar\hat{L}_{14}$	0	$i\hbar\hat{L}_{31}$	$-i\hbar\hat{L}_{23}$	0	0	0	$i\hbar\hat{L}_{45}$	0	0	$i\hbar\hat{L}_{46}$	$-i\hbar\hat{L}_{36}$	$-i\hbar\hat{L}_{35}$	0
$\hat{L}_{15}$	0	$i\hbar\hat{L}_{35}$	$-i\hbar\hat{L}_{25}$	$-i\hbar\hat{L}_{45}$	0	0	0	$-i\hbar\hat{L}_{12}$	$i\hbar\hat{L}_{31}$	$i\hbar\hat{L}_{56}$	0	0	0	$-i\hbar\hat{L}_{14}$	$i\hbar\hat{L}_{16}$
$\hat{L}_{25}$	$-i\hbar\hat{L}_{35}$	0	$i\hbar\hat{L}_{15}$	0	$-i\hbar\hat{L}_{45}$	0	$i\hbar\hat{L}_{12}$	0	$-i\hbar\hat{L}_{23}$	0	$i\hbar\hat{L}_{56}$	0	0	$-i\hbar\hat{L}_{24}$	$i\hbar\hat{L}_{26}$
$\hat{L}_{35}$	$i\hbar\hat{L}_{25}$	$-i\hbar\hat{L}_{15}$	0	0	0	$-i\hbar\hat{L}_{45}$	$-i\hbar\hat{L}_{31}$	$i\hbar\hat{L}_{23}$	0	0	0	$i\hbar\hat{L}_{56}$	0	$-i\hbar\hat{L}_{34}$	$i\hbar\hat{L}_{36}$
$\hat{L}_{16}$	0	$i\hbar\hat{L}_{36}$	$-i\hbar\hat{L}_{26}$	$-i\hbar\hat{L}_{46}$	0	0	$-i\hbar\hat{L}_{56}$	0	0	0	$-i\hbar\hat{L}_{12}$	$i\hbar\hat{L}_{31}$	$-i\hbar\hat{L}_{14}$	0	$-i\hbar\hat{L}_{15}$
$\hat{L}_{26}$	$-i\hbar\hat{L}_{36}$	0	$i\hbar\hat{L}_{16}$	0	$-i\hbar\hat{L}_{46}$	0	0	$-i\hbar\hat{L}_{56}$	0	$i\hbar\hat{L}_{12}$	0	$-i\hbar\hat{L}_{23}$	$-i\hbar\hat{L}_{24}$	0	$-i\hbar\hat{L}_{25}$
$\hat{L}_{36}$	$i\hbar\hat{L}_{26}$	$-i\hbar\hat{L}_{16}$	0	0	0	$-i\hbar\hat{L}_{46}$	0	0	$-i\hbar\hat{L}_{56}$	$-i\hbar\hat{L}_{31}$	$i\hbar\hat{L}_{23}$	0	$-i\hbar\hat{L}_{34}$	0	$-i\hbar\hat{L}_{35}$
$\hat{L}_{46}$	0	0	0	$i\hbar\hat{L}_{16}$	$i\hbar\hat{L}_{26}$	$i\hbar\hat{L}_{36}$	0	0	0	$i\hbar\hat{L}_{14}$	$i\hbar\hat{L}_{24}$	$i\hbar\hat{L}_{34}$	0	$-i\hbar\hat{L}_{56}$	$-i\hbar\hat{L}_{45}$
$\hat{L}_{45}$	0	0	0	$i\hbar\hat{L}_{15}$	$i\hbar\hat{L}_{25}$	$i\hbar\hat{L}_{35}$	$i\hbar\hat{L}_{14}$	$i\hbar\hat{L}_{24}$	$i\hbar\hat{L}_{34}$	0	0	0	$i\hbar\hat{L}_{56}$	0	$i\hbar\hat{L}_{46}$
$\hat{L}_{56}$	0	0	0	0	0	0	$-i\hbar\hat{L}_{16}$	$-i\hbar\hat{L}_{26}$	$-i\hbar\hat{L}_{36}$	$i\hbar\hat{L}_{15}$	$i\hbar\hat{L}_{25}$	$i\hbar\hat{L}_{35}$	$i\hbar\hat{L}_{45}$	$-i\hbar\hat{L}_{46}$	0

between the upper and lower blocks in the matrix. By combining the LRL operators $\hat{A}_i$ with $\hat{Q}_2$, we obtain the series $\hat{L}_{15}$, $\hat{L}_{25}$, $\hat{L}_{35}$, which we will denote by the symbol $\hat{B}_i$:

$$\left[\hat{Q}_2, \hat{A}_i\right] = i\hat{B}_i. \tag{12.30}$$

The three $\hat{B}_i$ operators are the components of a vector $\hat{\mathbf{B}}$ that is, in a sense, conjugate to the LRL vector. We get:

$$\hat{\mathbf{B}} = \frac{1}{2}\hat{\mathbf{R}}\hat{P}^2 - \hat{\mathbf{P}}\left(\hat{\mathbf{R}} \cdot \hat{\mathbf{P}}\right) + \frac{1}{2}\hat{\mathbf{R}}, \tag{12.31}$$

which is identical to the expression for $\hat{\mathbf{A}}$, except for a plus sign in the last term. Similarly, the upper three rows in the sixth column of Eq. (12.7) $\hat{L}_{16}$, $\hat{L}_{26}$, $\hat{L}_{36}$ are obtained by the commutators of the LRL vector with $\hat{Q}_1$:

$$\left[\hat{Q}_1, \hat{A}_i\right] = i\hat{\Gamma}_i. \tag{12.32}$$

The three operators $\hat{\Gamma}_i$ are the components of another vector, which Barut called $\hat{\mathbf{\Gamma}}$, and which appears to be proportional to the linear momentum:

$$\hat{\mathbf{\Gamma}} = R\hat{\mathbf{P}}. \tag{12.33}$$

In this way, fifteen operators are finally obtained, which may be verified to form an $\mathfrak{so}(4,2)$ algebra. The operators can be written in matrix form as

$$\mathbb{L} \Longleftrightarrow \begin{pmatrix} 0 & \hat{L}_3 & -\hat{L}_2 & \hat{A}_1 & \hat{B}_1 & \hat{\Gamma}_1 \\ & 0 & \hat{L}_1 & \hat{A}_2 & \hat{B}_2 & \hat{\Gamma}_2 \\ & & 0 & \hat{A}_3 & \hat{B}_3 & \hat{\Gamma}_3 \\ & & & 0 & \hat{Q}_2 & \hat{Q}_1 \\ & & & & 0 & \hat{Q}_3 \\ & & & & & 0 \end{pmatrix}. \tag{12.34}$$

The resulting commutation table is given in Table 12.2 and is clearly isomorphic to Table 12.1.

12.2 THE CARTAN-WEYL BASIS

Following our general strategy, we will now derive the *Cartan subalgebra* $\mathfrak{H}$ and corresponding *Weyl diagrams* of the $\mathfrak{so}(4,2)$ algebra. This will necessitate a change of basis—from the $\{\hat{\mathbf{L}}, \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{\Gamma}}, \hat{\mathbf{Q}}\}$ basis, introduced earlier, to the *Cartan-Weyl basis.* To this aim, we will follow the three steps from the first Scholium (Chapter 6), albeit in a slightly modified order. We will also be helped by the fact that the $\mathfrak{so}(4)$ algebra of the hydrogen atom (Chapter 9) forms a subalgebra of the spectrum-generating $\mathfrak{so}(4,2)$ algebra. The results in this section are easily verified on the basis of the elementary commutation rules in Table 12.2; most of the computations are therefore left to you.

12.2.1 Cartan subalgebra and Cartan generators

Let us start, as usual, with the identification of the maximal subset of commuting generators of the $\mathfrak{so}(4,2)$ algebra. This is easily identified in view of the simple commutation relationship in Eq. (12.9), which says that two operators commute only

Table 12.2 Commutation table for the angular momentum operators $\hat{L}_i$, LRL operators $\hat{A}_i$, radial momentum operators $\hat{Q}_i$, and vector operators $\hat{B}_i$ and $\hat{\Gamma}_i$ ($i = 1 \to 3$). The commutator $[\hat{X}_i, \hat{X}_j]$ is listed in the ith row and jth column. When compared with Table 12.1, it is clear that these fifteen operators form a *realization* of the $\mathfrak{so}(4,2)$ Lie algebra.

	$\hat{L}_1$	$\hat{L}_2$	$\hat{L}_3$	$\hat{A}_1$	$\hat{A}_2$	$\hat{A}_3$	$\hat{B}_1$	$\hat{B}_2$	$\hat{B}_3$	$\hat{\Gamma}_1$	$\hat{\Gamma}_2$	$\hat{\Gamma}_3$	$\hat{Q}_1$	$\hat{Q}_2$	$\hat{Q}_3$
$\hat{L}_1$	0	$i\hbar\hat{L}_3$	$-i\hbar\hat{L}_2$	0	$i\hbar\hat{A}_3$	$-i\hbar\hat{A}_2$	0	$i\hbar\hat{B}_3$	$-i\hbar\hat{B}_2$	0	$i\hbar\hat{\Gamma}_3$	$-i\hbar\hat{\Gamma}_2$	0	0	0
$\hat{L}_2$	$-i\hbar\hat{L}_3$	0	$i\hbar\hat{L}_1$	$-i\hbar\hat{A}_3$	0	$i\hbar\hat{A}_1$	$-i\hbar\hat{B}_3$	0	$i\hbar\hat{B}_1$	$-i\hbar\hat{\Gamma}_3$	0	$i\hbar\hat{\Gamma}_1$	0	0	0
$\hat{L}_3$	$i\hbar\hat{L}_2$	$-i\hbar\hat{L}_1$	0	$i\hbar\hat{A}_2$	$-i\hbar\hat{A}_1$	0	$i\hbar\hat{B}_2$	$-i\hbar\hat{B}_1$	0	$i\hbar\hat{\Gamma}_2$	$-i\hbar\hat{\Gamma}_1$	0	0	0	0
$\hat{A}_1$	0	$i\hbar\hat{A}_3$	$-i\hbar\hat{A}_2$	0	$i\hbar\hat{L}_3$	$-i\hbar\hat{L}_2$	$i\hbar\hat{Q}_2$	0	0	$i\hbar\hat{Q}_1$	0	0	$-i\hbar\hat{\Gamma}_1$	$-i\hbar\hat{B}_1$	0
$\hat{A}_2$	$-i\hbar\hat{A}_3$	0	$i\hbar\hat{A}_1$	$-i\hbar\hat{L}_3$	0	$i\hbar\hat{L}_1$	0	$i\hbar\hat{Q}_2$	0	0	$i\hbar\hat{Q}_1$	0	$-i\hbar\hat{\Gamma}_2$	$-i\hbar\hat{B}_2$	0
$\hat{A}_3$	$i\hbar\hat{A}_2$	$-i\hbar\hat{A}_1$	0	$i\hbar\hat{L}_2$	$-i\hbar\hat{L}_1$	0	0	0	$i\hbar\hat{Q}_2$	0	0	$i\hbar\hat{Q}_1$	$-i\hbar\hat{\Gamma}_3$	$-i\hbar\hat{B}_3$	0
$\hat{B}_1$	0	$i\hbar\hat{B}_3$	$-i\hbar\hat{B}_2$	$-i\hbar\hat{Q}_2$	0	0	0	$-i\hbar\hat{L}_3$	$i\hbar\hat{L}_2$	$i\hbar\hat{Q}_3$	0	0	0	$-i\hbar\hat{A}_1$	$i\hbar\hat{\Gamma}_1$
$\hat{B}_2$	$-i\hbar\hat{B}_3$	0	$i\hbar\hat{B}_1$	0	$-i\hbar\hat{Q}_2$	0	$i\hbar\hat{L}_3$	0	$-i\hbar\hat{L}_1$	0	$i\hbar\hat{Q}_3$	0	0	$-i\hbar\hat{A}_2$	$i\hbar\hat{\Gamma}_2$
$\hat{B}_3$	$i\hbar\hat{B}_2$	$-i\hbar\hat{B}_1$	0	0	0	$-i\hbar\hat{Q}_2$	$-i\hbar\hat{L}_2$	$i\hbar\hat{L}_1$	0	0	0	$i\hbar\hat{Q}_3$	0	$-i\hbar\hat{A}_3$	$i\hbar\hat{\Gamma}_3$
$\hat{\Gamma}_1$	0	$i\hbar\hat{\Gamma}_3$	$-i\hbar\hat{\Gamma}_2$	$-i\hbar\hat{Q}_1$	0	0	$-i\hbar\hat{Q}_3$	0	0	0	$-i\hbar\hat{L}_3$	$i\hbar\hat{L}_2$	$-i\hbar\hat{A}_1$	0	$-i\hbar\hat{B}_1$
$\hat{\Gamma}_2$	$-i\hbar\hat{\Gamma}_3$	0	$i\hbar\hat{\Gamma}_1$	0	$-i\hbar\hat{Q}_1$	0	0	$-i\hbar\hat{Q}_3$	0	$i\hbar\hat{L}_3$	0	$-i\hbar\hat{L}_1$	$-i\hbar\hat{A}_2$	0	$-i\hbar\hat{B}_2$
$\hat{\Gamma}_3$	$i\hbar\hat{\Gamma}_2$	$-i\hbar\hat{\Gamma}_1$	0	0	0	$-i\hbar\hat{Q}_1$	0	0	$-i\hbar\hat{Q}_3$	$-i\hbar\hat{L}_2$	$i\hbar\hat{L}_1$	0	$-i\hbar\hat{A}_3$	0	$-i\hbar\hat{B}_3$
$\hat{Q}_1$	0	0	0	$i\hbar\hat{\Gamma}_1$	$i\hbar\hat{\Gamma}_2$	$i\hbar\hat{\Gamma}_3$	0	0	0	$i\hbar\hat{A}_1$	$i\hbar\hat{A}_2$	$i\hbar\hat{A}_3$	0	$-i\hbar\hat{Q}_3$	$-i\hbar\hat{Q}_2$
$\hat{Q}_2$	0	0	0	$i\hbar\hat{B}_1$	$i\hbar\hat{B}_2$	$i\hbar\hat{B}_3$	$i\hbar\hat{A}_1$	$i\hbar\hat{A}_2$	$i\hbar\hat{A}_3$	0	0	0	$i\hbar\hat{Q}_3$	0	$i\hbar\hat{Q}_1$
$\hat{Q}_3$	0	0	0	0	0	0	$-i\hbar\hat{\Gamma}_1$	$-i\hbar\hat{\Gamma}_2$	$-i\hbar\hat{\Gamma}_3$	$i\hbar\hat{B}_1$	$i\hbar\hat{B}_2$	$i\hbar\hat{B}_3$	$i\hbar\hat{Q}_2$	$-i\hbar\hat{Q}_1$	0

if they do not have indices in common. In this way, we obtain three commuting operators, which we may choose as $\hat{L}_{12}$, $\hat{L}_{34}$, and $\hat{L}_{56}$ (i.e., $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$, respectively):

$$\left[\hat{L}_3, \hat{A}_3\right] = \left[\hat{L}_3, \hat{Q}_3\right] = \left[\hat{A}_3, \hat{Q}_3\right] = 0. \tag{12.35}$$

The triplet $\{\hat{L}_3, \hat{A}_3, \hat{Q}_3\}$ forms a basis for the *maximal Abelian Cartan subalgebra* $\mathfrak{H}$ of $\mathfrak{so}(4,2)$. The three operators $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$ are called *Cartan generators*, and the dimensionality of $\mathfrak{H}$ defines the *rank* of the $\mathfrak{so}(4,2)$ Lie algebra. As a consequence, all weight and root diagrams (introduced later) will be three-dimensional.

12.2.2 Casimir invariants

In view of Racah's theorem, and the fact that the $\mathfrak{so}(4,2)$ algebra is of rank 3, we expect the SO(4,2) group to exhibit three independent *Casimir invariants* $\hat{C}_\mu$ $(\mu = 1 \rightarrow 3)$ that commute with all the generators of the $\mathfrak{so}(4,2)$ algebra. The most important Casimir operator is the *trace operator* $\hat{C}_2$, which is a second-order combination of the invariants for the various subgroups:

$$\hat{C}_2 = \hat{L}^2 + \hat{A}^2 - \hat{B}^2 - \hat{\Gamma}^2 + \hat{Q}_3^2 - \hat{Q}_1^2 - \hat{Q}_2^2. \tag{12.36}$$

Here, $\hat{L}^2 + \hat{A}^2$ and $\hat{Q}_3^2 - \hat{Q}_1^2 - \hat{Q}_2^2$ are already familiar as the Casimir operators of the SO(4) and SO(2,1) groups, respectively. The remaining $\hat{B}^2$ and $\hat{\Gamma}^2$ refer to the squares of the $\hat{\mathbf{B}}$ and $\hat{\mathbf{\Gamma}}$ vectors, respectively. These terms are subtracted in $\hat{C}_2$ in line with their negative metric. For the hydrogenic realization, these scalar products can all be expressed in the $\hat{Q}$ operators, as follows:

$$\begin{aligned} \hat{L}^2 &= \hat{Q}_3^2 - \hat{Q}_1^2 - \hat{Q}_2^2; \\ \hat{A}^2 &= \hat{Q}_1^2 + \hat{Q}_2^2 - 1; \\ \hat{B}^2 &= \hat{Q}_3^2 - \hat{Q}_2^2 + 1; \\ \hat{\Gamma}^2 &= \hat{Q}_3^2 - \hat{Q}_1^2 + 1. \end{aligned} \tag{12.37}$$

The first expression was obtained in Chapter 11, Eq. (11.68). The expressions for $\hat{A}^2$, $\hat{B}^2$, and $\hat{\Gamma}^2$ are obtained in a similar way by direct calculation, starting from the radial expressions for these operators. Inserting these results into the expression for the Casimir operator yields an invariant that characterizes all the hydrogenic levels by a single integer:

$$\hat{C}_2 \left|nlm\right\rangle = -3 \left|nlm\right\rangle. \tag{12.38}$$

The remaining two Casimir operators are third- and fourth-order polynomials in the generators. Here are their definitions for completeness sake:

$$\hat{C}_3 = \frac{1}{48} \varepsilon_{abcdef} \hat{L}^{ab} \hat{L}^{cd} \hat{L}^{ef} = 0; \tag{12.39}$$

$$\hat{C}_4 = \hat{L}_{ab} \hat{L}^{bc} \hat{L}_{cd} \hat{L}^{da} = 0, \tag{12.40}$$

where ε is the completely antisymmetric tensor on six covariant indices with $\varepsilon_{123456} = 1$ and

$$\hat{L}^{ab} = g^{ac} g^{bd} \hat{L}_{cd} \quad \text{with } g^{ab} = g_{ab}, \tag{12.41}$$

where the Einstein summation convention has been used. The Casimir operators are used to label irreducible representations (or *multiplets*) of the SO(4,2) group. Because all of them are reduced to a constant value, there is only one unitary irreducible representation for the SO(4,2) group in the hydrogenic basis. That is to say, all the (ground and excited) states of the hydrogen atom $|nlm\rangle$ are contained in a single unirrep of $\mathfrak{so}(4,2)$.

12.2.3 Weyl generators

The remaining twelve non-Cartan operators can be recombined to yield commuting eigenoperators of the subalgebra $\mathfrak{H}$. The derivation of these *Weyl generators* will be presented in this section, along with the construction of a *root diagram* for the $\mathfrak{so}(4,2)$ algebra that hosts the Cartan and Weyl elements.

Weyl generators in the L_3A_3-plane

For each generator of the subalgebra $\mathfrak{H}$, there are six operators that do not have indices in common. For example, for $\hat{L}_{56}$ ($\hat{Q}_3$) the six off-diagonal elements are $\hat{L}_{12}$, $\hat{L}_{23}$, $\hat{L}_{31}$, $\hat{L}_{14}$, $\hat{L}_{24}$, and $\hat{L}_{34}$. In view of Eq. (12.9), these operators all commute with $\hat{L}_{56}$ and they will be positioned in the horizontal plane of the root diagram. They correspond to the components $\hat{L}_i$ and $\hat{A}_i$ of the angular momentum vector $\hat{\mathbf{L}}$ and the LRL vector $\hat{\mathbf{A}}$, and generate an $\mathfrak{so}(4)$ invariance algebra, as seen in Chapter 9. We can therefore take over our analysis of the $\mathfrak{so}(4)$ algebra from §9.5.3, which leads to a square root diagram in the horizontal L_3A_3-plane, as depicted in Figure 9.9.

Weyl generators in the L_3Q_3-plane

Let us see if a similar analysis can be pursued for the root diagrams in the vertical L_3Q_3- and A_3Q_3-planes. We start with a study of the L_3Q_3-plane. The six off-diagonal elements of $\hat{L}_{34}$ ($\hat{A}_3$) are $\hat{L}_{12}$, $\hat{L}_{15}$, $\hat{L}_{25}$, $\hat{L}_{16}$, $\hat{L}_{26}$, and $\hat{L}_{56}$. These operators correspond to the hydrogenic operators $\hat{L}_3$, $\hat{B}_1$, $\hat{B}_2$, $\hat{\Gamma}_1$, $\hat{\Gamma}_2$, and $\hat{Q}_3$ and are positioned in the L_3Q_3-plane. It will prove useful to rewrite them into the following linear combinations:

$$\begin{aligned}
\hat{M}_{11} &= \frac{1}{2}\left(\hat{B}_1+\hat{\Gamma}_2\right); & \hat{M}_{21} &= \frac{1}{2}\left(\hat{B}_1-\hat{\Gamma}_2\right); \\
\hat{M}_{12} &= \frac{1}{2}\left(\hat{B}_2-\hat{\Gamma}_1\right); & \hat{M}_{22} &= \frac{1}{2}\left(\hat{B}_2+\hat{\Gamma}_1\right); \\
\hat{M}_{13} &= \frac{1}{2}\left(\hat{L}_3+\hat{Q}_3\right); & \hat{M}_{23} &= \frac{1}{2}\left(\hat{L}_3-\hat{Q}_3\right).
\end{aligned} \tag{12.42}$$

It can then be shown that the components of the $\hat{\mathbf{M}}_1$ and $\hat{\mathbf{M}}_2$ operators mutually commute:

$$\left[\hat{M}_{1i},\hat{M}_{2j}\right]=0, \quad \forall i,j=1\rightarrow 3. \tag{12.43}$$

The components of $\hat{\mathbf{M}}_1$ commute among themselves to form an $\mathfrak{so}(2,1)$ algebra:

$$\left[\hat{M}_{11},\hat{M}_{12}\right]=-i\hbar\hat{M}_{13}, \quad \left[\hat{M}_{12},\hat{M}_{13}\right]=i\hbar\hat{M}_{11}, \quad \left[\hat{M}_{13},\hat{M}_{11}\right]=i\hbar\hat{M}_{12}, \tag{12.44}$$

and similarly for the components of $\hat{\mathbf{M}}_2$:

$$\left[\hat{M}_{21},\hat{M}_{22}\right]=-i\hbar\hat{M}_{23}, \quad \left[\hat{M}_{22},\hat{M}_{23}\right]=i\hbar\hat{M}_{21}, \quad \left[\hat{M}_{23},\hat{M}_{21}\right]=i\hbar\hat{M}_{22}. \tag{12.45}$$

We will denote these two algebras by $\mathfrak{so}(2,1)_1$ and $\mathfrak{so}(2,1)_2$, respectively, for further convenience. The commutation Table 12.3 shows that both algebras are completely

Table 12.3 Commutation table for the generators of the $\mathfrak{so}(2,2)$ Lie algebra in the basis $\{\hat{M}_{11}, \hat{M}_{12}, \hat{M}_{13}, \hat{M}_{21}, \hat{M}_{22}, \hat{M}_{23}\}$. The generators $\hat{\mathbf{M}}_1 = \left(\hat{M}_{11}, \hat{M}_{12}, \hat{M}_{13}\right)$ and $\hat{\mathbf{M}}_2 = \left(\hat{M}_{21}, \hat{M}_{22}, \hat{M}_{23}\right)$ each constitute an $\mathfrak{so}(2,1)$ subalgebra, denoted $\mathfrak{so}(2,1)_1$ and $\mathfrak{so}(2,1)_2$, respectively: $\mathfrak{so}(2,2) = \mathfrak{so}(2,1)_1 \oplus \mathfrak{so}(2,1)_2$. Both subalgebras are located in a square region, separated from one another by two squares of zeros.

	$\hat{M}_{11}$	$\hat{M}_{12}$	$\hat{M}_{13}$	$\hat{M}_{21}$	$\hat{M}_{22}$	$\hat{M}_{23}$
$\hat{M}_{11}$	0	$-i\hbar\hat{M}_{13}$	$-i\hbar\hat{M}_{12}$	0	0	0
$\hat{M}_{12}$	$i\hbar\hat{M}_{13}$	0	$i\hbar\hat{M}_{11}$	0	0	0
$\hat{M}_{13}$	$i\hbar\hat{M}_{12}$	$-i\hbar\hat{M}_{11}$	0	0	0	0
$\hat{M}_{21}$	0	0	0	0	$-i\hbar\hat{M}_{23}$	$-i\hbar\hat{M}_{22}$
$\hat{M}_{22}$	0	0	0	$i\hbar\hat{M}_{23}$	0	$i\hbar\hat{M}_{21}$
$\hat{M}_{23}$	0	0	0	$i\hbar\hat{M}_{22}$	$-i\hbar\hat{M}_{21}$	0

decoupled in view of Eq. (12.43). Notice, however, that instead of obtaining an $\mathfrak{so}(4)$ algebra as in Chapter 9 (which can be factorized into a direct sum of two $\mathfrak{so}(3)$ algebras), we now obtain an $\mathfrak{so}(2,2)$ Lie algebra, which is locally isomorphic to the direct sum of two $\mathfrak{so}(2,1)$ algebras:

$$\mathfrak{so}(2,2) = \mathfrak{so}(2,1)_1 \oplus \mathfrak{so}(2,1)_2. \tag{12.46}$$

Despite this difference, the $\mathfrak{so}(2,2)$ root diagram is similar to the one obtained in Figure 9.9. To see this, the two commuting operators $\hat{M}_{13}$ and $\hat{M}_{23}$ can be taken as a basis $\{\hat{M}_{13}, \hat{M}_{23}\}$ for the Cartan subalgebra of $\mathfrak{so}(2,2)$. The remaining generators $\hat{M}_{11}$, $\hat{M}_{12}$, $\hat{M}_{21}$, and $\hat{M}_{22}$ are then rearranged into the following linear combinations:

$$\begin{aligned} \hat{M}_{1+} &\equiv \hat{M}_{11} + i\hat{M}_{12}, \quad \hat{M}_{1-} \equiv \hat{M}_{11} - i\hat{M}_{12}, \\ \hat{M}_{2+} &\equiv \hat{M}_{21} + i\hat{M}_{22}, \quad \hat{M}_{2-} \equiv \hat{M}_{21} - i\hat{M}_{22}, \end{aligned} \tag{12.47}$$

which form a linearly independent set of *Weyl generators*. By virtue of their function as *ladder operators*, these Weyl elements act as *eigenoperators* of the Cartan generators:

$$\left[\hat{H}_i, \hat{E}_\alpha\right] = \alpha_i \hat{E}_\alpha, \quad \forall i = 1,2; \alpha = 1 \to 4, \tag{12.48}$$

where we have denoted the Cartan generators by the general symbol $\hat{H}_i$ $(i = 1,2)$ and the Weyl generators by the symbol $\hat{E}_\alpha$ $(\alpha = 1 \to 4)$. The *roots* α_1 and α_2 of every Weyl element $\hat{E}_\alpha$ can be read off from the commutation Table 12.4. They form the components of a two-dimensional *root vector* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ that can be positioned in a two-dimensional *weight space* formed by the $M_{11}M_{21}$-plane. The root vectors are

$$\begin{aligned} \boldsymbol{\alpha}\left(\hat{M}_{1+}\right) &= (\hbar, 0), \quad \boldsymbol{\alpha}\left(\hat{M}_{1-}\right) = (-\hbar, 0), \\ \boldsymbol{\alpha}\left(\hat{M}_{2+}\right) &= (0, \hbar), \quad \boldsymbol{\alpha}\left(\hat{M}_{2-}\right) = (0, -\hbar), \end{aligned} \tag{12.49}$$

and the corresponding *root diagram* is depicted in Figure 12.1. The mutually commuting Cartan generators $\hat{M}_{13}$ and $\hat{M}_{23}$ are at the origin of the diagram.

Weyl generators in the A_3Q_3-plane

The off-diagonal elements of $\hat{L}_{12}$ $(\hat{L}_3)$ are $\hat{L}_{34}$, $\hat{L}_{35}$, $\hat{L}_{36}$, $\hat{L}_{45}$, $\hat{L}_{46}$, and $\hat{L}_{56}$. They correspond to the operators $\hat{A}_3$, $\hat{B}_3$, $\hat{\Gamma}_3$, $\hat{Q}_2$, $\hat{Q}_1$, and $\hat{Q}_3$. Since they commute with $\hat{L}_3$,

Table 12.4 Commutation table for the generators of the $\mathfrak{so}(2,2)$ Lie algebra in the *Cartan-Weyl basis* $\{\hat{M}_{13}, \hat{M}_{1+}, \hat{M}_{1-}, \hat{M}_{23}, \hat{M}_{2+}, \hat{M}_{2-}\}$. The *Cartan subalgebra* $\mathfrak{H} = \{\hat{M}_{13}, \hat{M}_{23}\}$ forms a maximal Abelian subalgebra of $\mathfrak{so}(2,2)$, as indicated by the zeros in bold type in the table.

	$\hat{M}_{13}$	$\hat{M}_{1+}$	$\hat{M}_{1-}$	$\hat{M}_{23}$	$\hat{M}_{2+}$	$\hat{M}_{2-}$
$\hat{M}_{13}$	$\mathbf{0}$	$\hbar M_{1+}$	$-\hbar\hat{M}_{1-}$	$\mathbf{0}$	0	0
$\hat{M}_{1+}$	$-\hbar M_{1+}$	0	$-2\hbar\hat{M}_{13}$	0	0	0
$\hat{M}_{1-}$	$\hbar\hat{M}_{1-}$	$2\hbar\hat{M}_{13}$	0	0	0	0
$\hat{M}_{23}$	$\mathbf{0}$	0	0	$\mathbf{0}$	$\hbar\hat{M}_{2+}$	$-\hbar\hat{M}_{2-}$
$\hat{M}_{2+}$	0	0	0	$-\hbar\hat{M}_{2+}$	0	$-2\hbar\hat{M}_{23}$
$\hat{M}_{2-}$	0	0	0	$\hbar\hat{M}_{2-}$	$2\hbar\hat{M}_{23}$	0

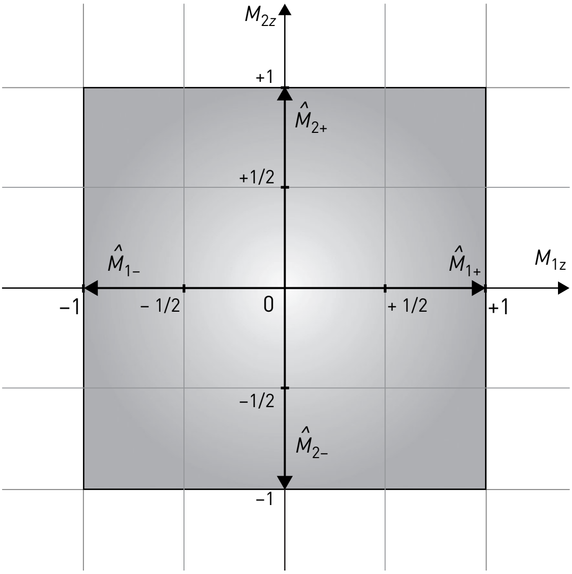


FIGURE 12.1 Root diagram of the $\mathfrak{so}(2,2)$ Lie algebra. The action of every Weyl generator is shown in the $M_{13}M_{23}$-plane. The Cartan generators $\hat{M}_{13}$ and $\hat{M}_{23}$ are positioned at the origin of the Weyl diagram.

they will all be positioned in the A_3Q_3-plane. If we rewrite them in the following linear combinations

$$\begin{aligned} \hat{N}_{11} &= \frac{1}{2}\left(\hat{\Gamma}_3 - \hat{Q}_2\right); \quad \hat{N}_{21} = \frac{1}{2}\left(\hat{B}_3 - \hat{Q}_1\right); \\ \hat{N}_{12} &= \frac{1}{2}\left(\hat{B}_3 + \hat{Q}_1\right); \quad \hat{N}_{22} = \frac{1}{2}\left(\hat{\Gamma}_3 + \hat{Q}_2\right); \\ \hat{N}_{13} &= \frac{1}{2}\left(\hat{A}_3 + \hat{Q}_3\right); \quad \hat{N}_{23} = \frac{1}{2}\left(\hat{A}_3 - \hat{Q}_3\right) \end{aligned} \tag{12.50}$$

a completely analogous analysis can be performed as in the last subsection. The components of $\hat{\mathbf{N}}_1$ and $\hat{\mathbf{N}}_2$ will thus mutually commute, and they will form a basis for two $\mathfrak{so}(2,1)$ algebras. This gives rise to an overall $\mathfrak{so}(2,2)$ Lie algebra, with a root diagram identical to the one in Figure 12.1, but where the $\hat{M}_{ij}$ operators are replaced by the $\hat{N}_{ij}$ operators.

12.2.4 SO(4,2) root diagram

Having analyzed the root structure of the $\mathfrak{so}(4)$ and $\mathfrak{so}(2,2)$ subalgebras of $\mathfrak{so}(4,2)$ in each of the orthogonal planes L_3A_3, L_3Q_3, and A_3Q_3, we are now in a position to bring together all the results and construct the *root diagram* for the $\mathfrak{so}(4,2)$ Lie algebra.

Notice that a total of eighteen operators have been introduced in the previous sections. The $\hat{J}_{j3}$, $\hat{M}_{j3}$, and $\hat{N}_{j3}$ (with $j=1$ or 2) are, however, not linearly independent, and they are therefore replaced by the operators $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$. This reduces the number of generators to fifteen as required by the $\mathfrak{so}(4,2)$ algebra.

The three commuting operators $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$ are then selected as generators of the Cartan subalgebra $\mathfrak{H} \subset \mathfrak{so}(4,2)$. They form a basis for a three-dimensional orthogonal frame and are at the origin of the root diagram, in view of Eq. (12.35). Along with the remaining twelve Weyl generators $\hat{J}_{j\pm}$, $\hat{M}_{j\pm}$, and $\hat{N}_{j\pm}$ (with $j=1$ or 2), they constitute the *Cartan-Weyl basis* for the $\mathfrak{so}(4,2)$ algebra.

Let the general symbols $\hat{H}_i$ ($i = 1 \rightarrow 3$) and $\hat{E}_\alpha$ ($\alpha = 1 \rightarrow 12$) denote the different Cartan and Weyl elements, respectively. To position the different $\hat{E}_\alpha$'s in the root diagram, the *roots* α_i of each Weyl element will have to be determined with respect to the three Cartan generators, according to the general Eq. (12.48). Notice that in the previous sections, the roots were taken with respect to $\hat{J}_{j3}$, $\hat{M}_{j3}$, and $\hat{N}_{j3}$ (with $j=1$ or 2), *not* with respect to $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$. The new *root vectors* thus obtained are

$$
\begin{aligned}
&\boldsymbol{\alpha}\left(\hat{J}_{1+}\right) = (+\hbar,+\hbar,0); && \boldsymbol{\alpha}\left(\hat{M}_{1+}\right) = (+\hbar,0,+\hbar); && \boldsymbol{\alpha}\left(\hat{N}_{1+}\right) = (0,+\hbar,+\hbar);\\
&\boldsymbol{\alpha}\left(\hat{J}_{1-}\right) = (-\hbar,-\hbar,0); && \boldsymbol{\alpha}\left(\hat{M}_{1-}\right) = (-\hbar,0,-\hbar); && \boldsymbol{\alpha}\left(\hat{N}_{1-}\right) = (0,-\hbar,-\hbar);\\
&\boldsymbol{\alpha}\left(\hat{J}_{2+}\right) = (+\hbar,-\hbar,0); && \boldsymbol{\alpha}\left(\hat{M}_{2+}\right) = (+\hbar,0,-\hbar); && \boldsymbol{\alpha}\left(\hat{N}_{2+}\right) = (0,+\hbar,-\hbar);\\
&\boldsymbol{\alpha}\left(\hat{J}_{2-}\right) = (-\hbar,+\hbar,0); && \boldsymbol{\alpha}\left(\hat{M}_{2-}\right) = (-\hbar,0,+\hbar); && \boldsymbol{\alpha}\left(\hat{N}_{2-}\right) = (0,-\hbar,+\hbar);\\
&\boldsymbol{\alpha}\left(\hat{L}_{3}\right) = (0,0,0); && \boldsymbol{\alpha}\left(\hat{A}_{3}\right) = (0,0,0); && \boldsymbol{\alpha}\left(\hat{Q}_{3}\right) = (0,0,0).
\end{aligned}
$$

It is evident that the four $\hat{J}_{j\pm}$ operators give rise to the characteristic SO(4) square, with vertex points (±1, ±1) in the L_3A_3-plane (in units of $\hbar$). Similar results are obtained for the $\hat{M}_{j\pm}$ and $\hat{N}_{j\pm}$ operators, each of which gives rise to a square of operators in the L_3Q_3-and A_3Q_3-planes, respectively. The presentation of the fifteen generators in the root diagram of the $\mathfrak{so}(4,2)$ algebra thus consists of three generators at the origin and three squares in three perpendicular coordinate planes. The resulting figure is a *cuboctahedron*, as shown in Figure 12.2, with the Cartan generators located at the center and each Weyl generator pointing at one of the twelve vertices.

The cuboctahedron belongs to the thirteen semiregular *Archimedean solids.*[5] It has twelve vertices, fourteen faces and twenty-four edges, and is obtained in one of two

---

[5] These solids are composed of at least two different types of regular polygons. In this sense, *Archimedean solids* are different from the five *Platonic solids,* which are composed of only one type of polygon.

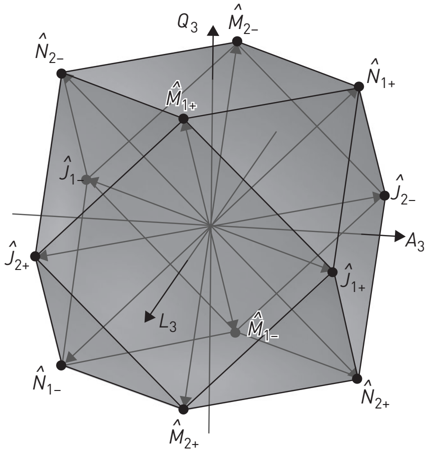


FIGURE 12.2 Root diagram of the $\mathfrak{so}(4,2)$ Lie algebra. The action of every Weyl generator is shown in the L_3–A_3–Q_3 space. The Cartan generators $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$ are positioned at the origin of the Weyl diagram.

FIGURE 12.3 M. C. Escher's 1947 mezzotint, titled *Crystal*. The intersection of the cube with the octahedron is a cuboctahedron. Reprinted with permission of the Escher Foundation, Den Haag.

ways: either by truncating a cube or by truncating an octahedron (hence, also the contraction cuboctahedron). That is, if the eight vertices of a cube are cut off by planes that bisect the edges meeting at each vertex, one obtains a polyhedron with eight equilateral triangular faces and six square faces. Alternatively, the six vertices of an octahedron can be cut off in a manner analogous to the truncated cube. A cuboctahedron is thus looked upon as the intersection of a cube with an octahedron, as shown in Escher's lithograph *Crystal* (Figure 12.3).

12.2.5 SO(4,2) weight diagrams

In a final step, a *weight diagram* can be constructed for the SO(4,2) group. The weight space is defined by the three Cartan generators $\hat{L}_3$, $\hat{A}_3$, and $\hat{Q}_3$, which function once again as a basis for a three-dimensional orthogonal frame.

The horizontal planes formed by the first two operators $\hat{L}_3$ and $\hat{A}_3$ host the different $\mathfrak{so}(4)$ manifolds. In Chapter 9, these manifolds were based on the operators $\hat{J}_{13}$ and $\hat{J}_{23}$, which in the current diagram correspond to the diagonal directions in view of their defining property:

$$\hat{J}_{13} = \frac{1}{2}\left(\hat{L}_3 + \hat{A}_3\right); \quad \hat{J}_{23} = \frac{1}{2}\left(\hat{L}_3 - \hat{A}_3\right). \tag{12.51}$$

The vertical direction, formed by the eigenvalues of the operator $\hat{Q}_3$, adds the radial ladder operator to the manifold. In this way, all the bound states of hydrogen are reunited in one single infinite-dimensional degeneracy space of SO(4,2). This space has been denoted by Carl E. Wulfman as the *baruton*, in honor of Barut.[6] Filling all the states with electrons leads to all the elements of the periodic system, and in this way the *baruton* can be looked upon as a massive particle that covers the full Periodic Table (see Chapter 13). The graphical representation shown in Figure 12.4 resembles a *square pyramid* placed upside down. This construction is also referred to as an SO(4,2) *tower*.

For a further study of the Periodic Table, the standard SO(4,2) tower is not really adequate because the SO(4) "floors" on which it is built are not really appropriate for the heavier elements. An alternative presentation makes use of the $|nlm\rangle$ orbitals, which are arranged in a *triangular* tower, as indicated in Figure 12.5. A given floor is characterized by the principal quantum number *n*. The horizontal stripes correspond to the different *l* subshells, and the dots are the individual *m* components. The allowed moves on this tower will be discussed in the next section.

The collection of all bound quantum levels in a single degenerate SO(4,2) manifold may be attributed to the *scaling* of the radial distance. The scaling parameter *a* is proportional to the square root of the energy. The ground level, which has the largest binding energy, is characterized by the smallest mean radius but is expanded most

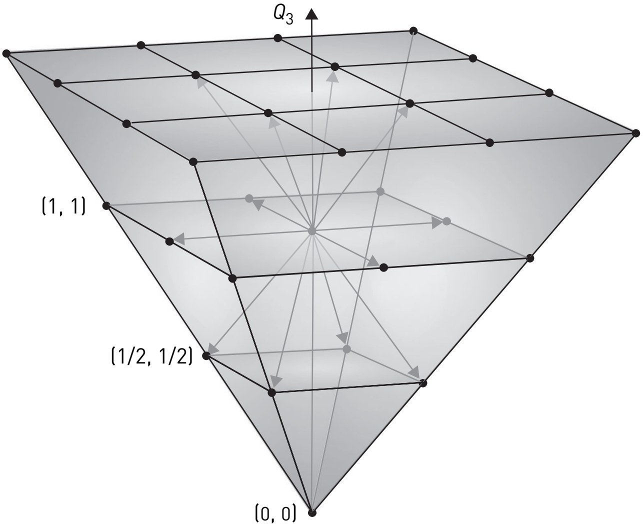


FIGURE 12.4 Weight diagram of the $\mathfrak{so}(4,2)$ Lie algebra, depicting the $(0,0)$, $(1/2,1/2)$, $(1,1)$ and $(3/2,3/2)$ manifolds from Figure 9.11. The action of the twelve Weyl generators is shown in the SO(4,2) *tower* by arrows pointing from the central state in the $(1,1)$ manifold.

---

[6] C. E. Wulfman. "Dynamical Groups in Atomic and Molecular Physics," p. 382.

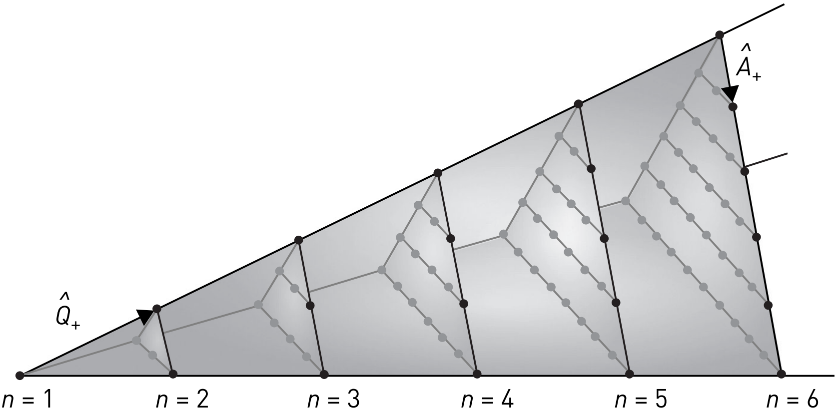


FIGURE 12.5 Triangular weight diagram of the $\mathfrak{so}(4,2)$ Lie algebra. The shaded triangles represent different SO(4) multiplets (for $n = 1 \rightarrow 6$). The horizontal lines correspond to different l-values, and the dots are the individual m components. The $\hat{Q}_\pm$ operators connect states of different n but fixed l, as indicated for the upper states with $n = 1$ or 2, $l = 0$, and $m = 0$. The $\hat{A}_\pm$ operators, on the other hand, link states of different l within a given n-multiplet. The $\hat{L}_\pm$ operators, finally, allow us to move along a horizontal line, keeping n and l fixed and varying m. [Adapted from B. G. Adams, J. Čížek, and J. Paldus. "Representation Theory of SO(4, 2) for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods." *International Journal of Quantum Chemistry* 21.1 (1982), pp. 153–171, p. 167.]

because it has the largest scaling factor. In this way the gradual extension of the bound levels with the higher principal quantum number n is compensated exactly by the rescaling, and all levels end up at the same quantum state.

12.3 QUANTUM ALCHEMY

In SO(4,2), the infinite number of hydrogen orbitals form the galleries of a universal library. In this section, we will construct the staircases and air shafts that allow us to move from one orbital to another. The term *quantum alchemy* refers to the transmutation of chemical elements that results from changing valence orbitals.[7]

12.3.1 Raising and lowering *m*

As discussed in Chapter 5, the ladder operators of angular momentum theory are designed to raise or lower the m quantum number. This may be denoted as follows:

$$\hat{L}_+ |nlm\rangle = \omega^l_{+m} |nl(m+1)\rangle; \tag{12.52}$$

$$\hat{L}_- |nlm\rangle = \omega^l_{-m} |nl(m-1)\rangle. \tag{12.53}$$

For our purposes, let us go one step further and identify the unknown coefficients $\omega^l_{\pm m}$ in these equations. This can be achieved by the algebraic properties of the

---

[7] S. M. Blinder, "Quantum Alchemy: Transmutation of Atomic Orbitals." *Journal of Chemical Education* 78.3 (2001). pp. 391–394.

operators themselves, in conjunction with normalization. The Casimir operator $\hat{L}^2$ may be written in four different forms:

$$\begin{aligned}\hat{L}^2 &= \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \\ &= \frac{1}{2}\left(\hat{L}_+\hat{L}_- + \hat{L}_-\hat{L}_+\right) + \hat{L}_z^2 \\ &= \hat{L}_+\hat{L}_- + \hat{L}_z^2 - \hbar\hat{L}_z \\ &= \hat{L}_-\hat{L}_+ + \hat{L}_z^2 + \hbar\hat{L}_z.\end{aligned} \tag{12.54}$$

The total angular momentum is given by the diagonal matrix element of $\hat{L}^2$. This can be worked out with the help of the previous expressions as

$$\begin{aligned}\langle nlm|\,\hat{L}^2\,|nlm\rangle &= \langle nlm|\,\hat{L}_-\hat{L}_+ + \hat{L}_z^2 + \hbar\hat{L}_z\,|nlm\rangle \\ &= \langle nlm|\,\hat{L}_-\hat{L}_+\,|nlm\rangle + \hbar^2\left(m^2+m\right) \\ &= \left\langle \hat{L}_+ nlm|\hat{L}_+ nlm\right\rangle + \hbar^2\left(m^2+m\right) \\ &= \left|\omega_{+m}^l\right|^2 + \hbar^2\left(m^2+m\right) \\ &= \hbar^2 l(l+1).\end{aligned}$$

Hence, we have

$$\left|\omega_{+m}^l\right|^2 = \hbar^2\left[l(l+1) - m(m+1)\right], \tag{12.55}$$

and similarly for the lowering operator:

$$\left|\omega_{-m}^l\right|^2 = \hbar^2\left[l(l+1) - m(m-1)\right]. \tag{12.56}$$

Since absolute values should be real positive or zero, these equations imply requirements on l and m that are precisely the ones we derived in the previous chapter for SO(3) representations (with $l = j$); hence,

$$(l \mp m)(l \pm m + 1) \geq 0. \tag{12.57}$$

For the actual values of the coefficients, we have to add a phase. Here, the choice of phase is a matter of selecting the relative phase between the hydrogenic functions. We follow the convention of B. G. Adams and colleagues, who simply took the coefficients to be positive.[8] Hence, we get

$$\omega_{\pm m}^l = \hbar\sqrt{(l \mp m)(l \pm m + 1)}. \tag{12.58}$$

---

[8] B. G. Adams, J. Čížek, and J. Paldus. "Representation Theory of SO(4, 2) for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods." *International Journal of Quantum Chemistry* 21.1 (1982), pp. 153–171 and B. G. Adams, J. Čížek, and J. Paldus. "Lie Algebraic Methods and Their Application to Simple Quantum Systems." In: *Advances in Quantum Chemistry*, vol. 19. Ed. Per-Olov Löwdin. San Diego: Academic Press, 1988, pp. 1–84, reproduced in B. G. Adams, J. Čížek, and J. Paldus. "Lie Algebraic Methods and Their Application to Simple Quantum Systems." In: *Dynamical Groups and Spectrum Generating Algebras*, vol. 1. Ed. Y. Ne'eman, A. Bohm, and A. O. Barut. Singapore: World Scientific Publishers, 1988, pp. 103–208. See also B. G. Adams. *Algebraic Approach to Simple Quantum Systems*. Berlin: Springer-Verlag, 1994, and J. Čížek and J. Paldus. "An Algebraic Approach to Bound States of Simple One-Electron Systems." *International Journal of Quantum Chemistry* 12.5 (1977), pp. 875–896.

By repeated application of the $\hat{L}_\pm$ operators, we are able to generate the entire ladder of states, with m varying from $-l$ to $+l$. Note that the raising operator coefficient vanishes for $m = l$; it is thus impossible to raise m above that threshold. Similarly, the lowering operator coefficient vanishes when m has reached the $m = -l$ lower boundary.

12.3.2 Raising and lowering *n*

In view of the structural correspondence between the compact SO(3) and noncompact SO(2,1) groups, the application of raising and lowering operators for the principal quantum number n is very similar to the ladder operations in m. The creation and annihilation operators are defined as in Chapter 11:

$$\hat{Q}_+ = \hat{Q}_1 + i\hat{Q}_2; \tag{12.59}$$

$$\hat{Q}_- = \hat{Q}_1 - i\hat{Q}_2. \tag{12.60}$$

These operators raise and lower the principal quantum number n, respectively, while keeping the lm labels fixed. This can be expressed as follows:

$$\hat{Q}_+ \left|nlm\right\rangle = \omega_l^{+n} \left|(n+1)lm\right\rangle; \tag{12.61}$$

$$\hat{Q}_- \left|nlm\right\rangle = \omega_l^{-n} \left|(n-1)lm\right\rangle. \tag{12.62}$$

The $\omega_l^{\pm n}$ coefficients can be found in a way that is entirely similar to the derivation of the $\omega_{\pm m}^{l}$ coefficients. First, the total momentum Q^2 is analyzed in its components, as was shown in Chapter 11, and is resumed here:

$$\begin{aligned}
\hat{Q}^2 &= \hat{Q}_3^2 - \hat{Q}_1^2 - \hat{Q}_2^2 \\
&= -\frac{1}{2}\left(\hat{Q}_+\hat{Q}_- + \hat{Q}_-\hat{Q}_+\right) + \hat{Q}_3^2 \\
&= -\hat{Q}_+\hat{Q}_- + \hat{Q}_3^2 - \hbar\hat{Q}_3 \\
&= -\hat{Q}_-\hat{Q}_+ + \hat{Q}_3^2 + \hbar\hat{Q}_3.
\end{aligned} \tag{12.63}$$

We then evaluate the matrix element of $\hat{Q}^2$, making use of the Hermiticity of the raising and lowering operators, and keeping in mind that the total Q^2 momentum is equal to the angular momentum:

$$\begin{aligned}
\left\langle nlm\right| \hat{Q}^2 \left|nlm\right\rangle &= \left\langle nlm\right| -\hat{Q}_-\hat{Q}_+ + \hat{Q}_3^2 + \hbar\hat{Q}_3 \left|nlm\right\rangle \\
&= -\left\langle nlm\right| \hat{Q}_-\hat{Q}_+ \left|nlm\right\rangle + \hbar^2\left(n^2+n\right) \\
&= -\left\langle \hat{Q}_+ nlm \middle| \hat{Q}_+ nlm \right\rangle + \hbar^2\left(n^2+n\right) \\
&= -\left|\omega_l^{+n}\right|^2 + \hbar^2\left(n^2+n\right) \\
&= \hbar^2 l(l+1).
\end{aligned} \tag{12.64}$$

Hence, we have

$$\left|\omega_l^{+n}\right|^2 = \hbar^2\left[n(n+1) - l(l+1)\right] \tag{12.65}$$

and, similarly for the lowering operator,

$$\left|\omega_l^{-n}\right|^2 = \hbar^2\left[n(n-1) - l(l+1)\right]. \tag{12.66}$$

These absolute values should be real positive or zero; hence, they imply requirements on l and n that are precisely the ones we derived in the previous chapter for SO(2,1) representations (with $l = j$ and $n = m$). Hence,

$$(l \mp n)(l \pm n + 1) \leq 0. \tag{12.67}$$

Again, only the absolute values of these coefficients are fixed by this expression, and the phase can be chosen freely. We have adopted here the simplest positive square root, following Adams[9] and colleagues, which leads to

$$\omega_l^{\pm n} = \hbar\sqrt{(n \mp l)(n \pm l \pm 1)}. \tag{12.68}$$

The lowest level one can reach by consecutive application of the lowering operator $\hat{Q}_-$ is the 1s ground level (with $n = 1$, $l = 0$), at which point the ω_l^{-n} coefficient vanishes. On the other hand, because the maximal l-value of a given level is equal to $n - 1$, the ω_l^{+n} coefficient is never zero, and thus there is no upper bound to n. We can continue to climb up in energy by applying the raising operator $\hat{Q}_+$ *ad infinitum.*

12.3.3 Raising and lowering *l*

The raising and lowering of the orbital quantum number l is less straightforward. Note that SO(4,2) contains the angular momentum operator $\hat{L}^2$ that yields $l(l+1)$, but is lacking an operator that yields l itself.[10] This makes a difference with the treatment of the m and n quantum numbers in the previous sections. The action of the LRL operators in the $|nlm\rangle$ manifold was derived in Chapter 9. The result is

$$\hat{A}^2|nlm\rangle = \hbar^2\left(n^2 - l(l+1) - 1\right)|nlm\rangle. \tag{12.69}$$

Raising and lowering operators are now defined in the same way as for angular momentum: $\hat{A}_\pm = \hat{A}_x \pm i\hat{A}_y$. In Table 12.5, we see the commutation relations for these operators. From these results, we may rewrite the squared operator as follows:

$$\begin{aligned}\hat{A}^2 &= \frac{1}{2}\left(\hat{A}_+\hat{A}_- + \hat{A}_-\hat{A}_+\right) + \hat{A}_z^2 \\ &= \hat{A}_-\hat{A}_+ + \hat{A}_z^2 + \hbar\hat{L}_z \\ &= \hat{A}_+\hat{A}_- + \hat{A}_z^2 - \hbar\hat{L}_z.\end{aligned} \tag{12.70}$$

Another useful relation is the orthogonality of the LRL and angular momentum vectors: $\hat{\mathbf{A}} \cdot \hat{\mathbf{L}} = 0$. This yields

$$\hat{\mathbf{A}} \cdot \hat{\mathbf{L}} = \frac{1}{2}\left(\hat{A}_+\hat{L}_- + \hat{A}_-\hat{L}_+\right) + \hat{A}_z\hat{L}_z = 0. \tag{12.71}$$

Since, in this expression, the lowering operator $\hat{A}_-$ is associated with the raising operator $\hat{L}_+$, it is clear that $\hat{A}_-$ lowers m by 1 unit, and vice versa for the raising operator $\hat{A}_+$, which must increase m by 1 unit. This is obvious from the commutation relations of $\hat{A}_\pm$ with $\hat{L}_z$ in Table 12.5. However, it is not so that they also lower or increase l accordingly.

---

[9] B. G. Adams, J. Čížek, and J. Paldus. "Representation Theory of SO(4, 2) for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods." 1982.

[10] In Chapter 14, we will introduce an extra operator $\hat{S}$ for this purpose. However, this operator does not belong to SO(4,2).

Table 12.5 Commutation table for the $\hat{\mathbf{L}}$ and $\hat{\mathbf{A}}$ operators in ladder format.

	$\hat{L}_+$	$\hat{L}_-$	$\hat{L}_z$	$\hat{A}_+$	$\hat{A}_-$	$\hat{A}_z$
$\hat{L}_+$	0	$2\hbar\hat{L}_z$	$-\hbar\hat{L}_+$	0	$2\hbar\hat{A}_z$	$-\hbar\hat{A}_+$
$\hat{L}_-$	$-2\hbar\hat{L}_z$	0	$\hbar\hat{L}_-$	$-2\hbar\hat{A}_z$	0	$\hbar\hat{A}_-$
$\hat{L}_z$	$\hbar\hat{L}_+$	$-\hbar\hat{L}_-$	0	$\hbar\hat{A}_+$	$-\hbar\hat{A}_-$	0
$\hat{A}_+$	0	$2\hbar\hat{A}_z$	$-\hbar\hat{A}_+$	0	$2\hbar\hat{L}_z$	$-\hbar\hat{L}_+$
$\hat{A}_-$	$-2\hbar\hat{A}_z$	0	$\hbar\hat{A}_-$	$-2\hbar\hat{L}_z$	0	$\hbar\hat{L}_-$
$\hat{A}_z$	$\hbar\hat{A}_+$	$-\hbar\hat{A}_-$	0	$\hbar\hat{L}_+$	$-\hbar\hat{L}_-$	0

The required l ladder operators must combine operators that have the same effect on m. Let us for the moment construct ladder operators that change l but leave m unchanged. As we have seen, there are three such operators: $\hat{A}_-\hat{L}_+$, $\hat{A}_+\hat{L}_-$, and $\hat{A}_z$. In view of Eq. (12.71), these three operators are not linearly independent. However, we can easily project two orthogonal operators, because the difference, $\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-$, is linearly independent of $\hat{A}_z$. Hence, the two operators $\{1/2\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right), \hat{A}_z\}$ form an orthogonal operator basis. The operator that will raise l by 1 unit will be denoted as $\hat{\Delta}_l^+$ and is expressed as the linear combination in the resulting operator basis:

$$\hat{\Delta}_l^+ = \frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + x\hbar\hat{A}_z. \tag{12.72}$$

Here, x is an unknown coefficient to be determined from the commutator of $\hat{L}^2$ and Δ_l^+. To raise l by 1 unit, this commutator must be equal to

$$\left[\hat{L}^2, \hat{\Delta}_l^+\right] = 2\hbar^2(l+1)\,\hat{\Delta}_l^+. \tag{12.73}$$

This requirement guarantees that the result of applying the raising operator to an l wave function is indeed an $(l+1)$ wave function, since

$$\begin{aligned} \hat{L}^2\hat{\Delta}_l^+|nlm\rangle &= \hat{\Delta}_l^+\left(\hat{L}^2 + \hbar^2(2l+2)\right)|nlm\rangle \\ &= \hbar^2\left(l(l+1) + 2(l+1)\right)\hat{\Delta}_l^+|nlm\rangle \\ &= \hbar^2(l+1)(l+2)\,\hat{\Delta}_l^+|nlm\rangle. \end{aligned} \tag{12.74}$$

The commutator can now be calculated using the results in Table 12.5. The $\hat{L}^2$ operator can be expanded as $\hat{L}_z^2 + \hbar\hat{L}_z + \hat{L}_-\hat{L}_+$. Since the operators under consideration commute with $\hat{L}_z$, as they leave m unchanged, we only have to consider commutators with $\hat{L}_-\hat{L}_+$. We have for the component operators:

$$\begin{aligned} \left[\hat{L}_-\hat{L}_+, \hat{A}_-\hat{L}_+\right] &= -2\hbar\hat{A}_-\hat{L}_+\hat{L}_z + 2\hbar\hat{A}_z\hat{L}_-\hat{L}_+; \\ \left[\hat{L}_-\hat{L}_+, \hat{A}_+\hat{L}_-\right] &= 2\hbar\hat{A}_+\hat{L}_-\hat{L}_z - 2\hbar\hat{A}_z\hat{L}_-\hat{L}_+ - 4\hbar^2\hat{A}_z\hat{L}_z; \\ \left[\hat{L}_-\hat{L}_+, \hat{A}_z\right] &= \hbar\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + 2\hbar^2\hat{A}_z. \end{aligned} \tag{12.75}$$

The commutator requirement of Eq. (12.74) can then be expressed as

$$\begin{aligned}\left[\hat{L}^2, \hat{\Delta}_l^+\right] &= 2\hbar\hat{A}_z\hat{L}_z^2 + 2\hbar\hat{A}_z\hat{L}_-\hat{L}_+ + 2\hbar^2\hat{A}_z\hat{L}_z \\ &\quad + x\hbar^2\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + 2x\hbar^3\hat{A}_z \\ &= 2\hbar^2(l+1)\left(\frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + x\hbar\hat{A}_z\right).\end{aligned} \tag{12.76}$$

Since the l raising operator was designed specifically for its action on the $|nlm\rangle$ hydrogenic ket, we can replace the angular momenta in this expression by their eigenvalues:

$$\begin{aligned}\hat{L}_z &\to \hbar m; \\ \hat{L}_-\hat{L}_+ &\to \hbar^2(l-m)(l+m+1).\end{aligned} \tag{12.77}$$

The commutator equation then is turned into a linear equation in the operator basis. Since the basis is orthogonal, this equation splits into two separate requirements: one on the coefficients preceding $1/2\left(\hat{A}_-\hat{L}_- - \hat{A}_+\hat{L}_-\right)$ and one on the coefficients preceding $\hat{A}_z$. These are

$$\begin{aligned}2x &= 2(l+1); \\ 2l^2 + 2l + 2x &= 2x(l+1).\end{aligned} \tag{12.78}$$

These equations are consistent and have as unique solution: $x = l + 1$. Hence, the raising operator reads

$$\hat{\Delta}_l^+ = \frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + (l+1)\hbar\hat{A}_z. \tag{12.79}$$

Note that this is not an ordinary Lie operator because it contains the l quantum number as a parameter. To apply this operator to a given function, the l characteristic number of this function should already be known in advance! So, this treatment is limited to the hydrogen eigenspace, with components $|nlm\rangle$.

To construct the inverse operator, which lowers l by 1 unit, we apply Hermitian conjugation to the raising operator, yielding

$$\hat{\Delta}_l^- = \frac{1}{2}\left(\hat{L}_-\hat{A}_+ - \hat{L}_+\hat{A}_-\right) + y\hbar\hat{A}_z. \tag{12.80}$$

Note that this conjugation cannot be fully executed, because it is not defined what will become of the numerical parameter x of the previous expression.[11] So, we keep an unknown coefficient, which is again to be determined from the commutation relation with $\hat{L}^2$. Indeed, for Δ_l^- to lower l by 1 unit, its commutator with $\hat{L}^2$ should be as follows:

$$\left[\hat{L}^2, \hat{\Delta}_l^-\right] = -2l\hbar^2\hat{\Delta}_l^-. \tag{12.81}$$

This implies that $\hat{\Delta}_l^-$ lowers l by one:

$$\begin{aligned}\hat{L}^2\hat{\Delta}_l^-|nlm\rangle &= \hat{\Delta}_l^-\left[\hat{L}^2 - 2\hbar^2 l\right]|nlm\rangle \\ &= \hbar^2\left[l(l+1) - 2l\right]\hat{\Delta}_l^-|nlm\rangle \\ &= \hbar^2 l(l-1)\,\hat{\Delta}_l^-|nlm\rangle.\end{aligned} \tag{12.82}$$

---

[11] A rigorous commutation conjugation requires the additional $\hat{S}$ operator, which is outside SO(4,2). It will be discussed in Chapter 14.

To work out the commutators, we note that the lowering operator can first be rewritten as

$$\hat{\Delta}_l^- = -\frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + (y-2)\,\hbar\hat{A}_z. \tag{12.83}$$

Making use of Eq. (12.75), we then find

$$\begin{aligned}\left[\hat{L}^2, \hat{\Delta}_l^-\right] &= -2\hbar\hat{A}_z\hat{L}_z^2 - 2\hbar\hat{A}_z\hat{L}_-\hat{L}_+ - 2\hbar^2\hat{A}_z\hat{L}_z \\ &\quad + (y-2)\,\hbar^2\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + 2\,(y-2)\,\hbar^3\hat{A}_z \\ &= -2l\hbar^2\left(-\frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + (y-2)\,\hbar\hat{A}_z\right).\end{aligned} \tag{12.84}$$

Again, this expression has a unique solution for the y-parameter, given by

$$y = l + 2. \tag{12.85}$$

The ladder operators for l are thus found to be

$$\hat{\Delta}_l^+ = \frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + (l+1)\hbar\hat{A}_z; \tag{12.86}$$

$$\hat{\Delta}_l^- = -\frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + l\hbar\hat{A}_z. \tag{12.87}$$

The action of the $\hat{\Delta}_l^{\pm}$ operators on the triplet of quantum numbers (nlm) introduces two unknown coefficients, c_{nlm}^+ and c_{nlm}^-, according to

$$\hat{\Delta}_l^+\,|nlm\rangle = c_{nlm}^+\,|n(l+1)\,m\rangle; \tag{12.88}$$

$$\hat{\Delta}_l^-\,|nlm\rangle = c_{nlm}^-\,|n(l-1)\,m\rangle. \tag{12.89}$$

The rest of this section will be devoted to finding these two unknown coefficients. To this aim, let us start by evaluating how the component operators $\hat{A}_z, \hat{A}_+$, and $\hat{A}_-$ act on the wave function. The action of $\hat{A}_z$ can be found easily. We simply have to add the two $\hat{\Delta}_l^{\pm}$ operators, as is shown here:

$$\hbar\hat{A}_z\,|nlm\rangle = \frac{c_{nlm}^+}{2l+1}\,|n\,(l+1)\,m\rangle + \frac{c_{nlm}^-}{2l+1}\,|n\,(l-1)\,m\rangle. \tag{12.90}$$

To find out how $\hat{A}_\pm$ act on the wave function, we use a sum and difference equation based respectively on Eq. (12.71) and the weighted difference of the $\hat{\Delta}_l^{\pm}$:

$$\begin{aligned}\frac{1}{2}\left(\hat{A}_-\hat{L}_+ + \hat{A}_+\hat{L}_-\right) &= -\hat{A}_z\hat{L}_z = -\frac{1}{\hbar\,(2l+1)}\hat{\Delta}_l^+\hat{L}_z - \frac{1}{\hbar\,(2l+1)}\hat{\Delta}_l^-\hat{L}_z; \\ \frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) &= \frac{l}{2l+1}\hat{\Delta}_l^+ - \frac{l+1}{2l+1}\hat{\Delta}_l^-.\end{aligned} \tag{12.91}$$

Addition of the results yields an expression for the $\hat{A}_-\hat{L}_+$ operator:

$$\hat{A}_-\hat{L}_+\,|nlm\rangle = \frac{c_{nlm}^+\,(l-m)}{2l+1}\,|n(l+1)m\rangle - \frac{c_{nlm}^-\,(l+m+1)}{2l+1}\,|n(l-1)m\rangle. \tag{12.92}$$

Now we apply $\hat{L}_-$ to both sides of this equation. Since the two lowering operators commute, $\left[\hat{L}_-, \hat{A}_-\right] = 0$, we can easily derive the desired action of $\hat{A}_-$:

$$\begin{aligned}\hat{A}_-\,|nlm\rangle &= \frac{c_{nlm}^+}{2l+1}\sqrt{\frac{l-m+2}{l+m+1}}\,|n(l+1)(m-1)\rangle \\ &\quad - \frac{c_{nlm}^-}{2l+1}\sqrt{\frac{l+m-1}{l-m}}\,|n(l-1)(m-1)\rangle.\end{aligned} \tag{12.93}$$

An entirely similar procedure also yields the $\hat{A}_+$ operator:

$$\hat{A}_+\hat{L}_- |nlm\rangle = -\frac{c^+_{nlm}(l+m)}{2l+1}|n(l+1)\,m\rangle + \frac{c^-_{nlm}(l-m+1)}{2l+1}|n(l-1)\,m\rangle . \quad (12.94)$$

Operating now left and right with the angular momentum raising operator $\hat{L}_+$, and making use of the fact that the $\hat{A}_+$ and $\hat{L}_+$ raising operators commute, we obtain

$$\begin{aligned} A_+|nlm\rangle = &-\frac{c^+_{nlm}}{2l+1}\sqrt{\frac{l+m+2}{l-m+1}}\,|n(l+1)(m+1)\rangle \\ &+\frac{c^-_{nlm}}{2l+1}\sqrt{\frac{l-m-1}{l+m}}\,|n(l-1)(m+1)\rangle. \end{aligned} \quad (12.95)$$

Armed with these expressions for the three components of $\hat{\mathbf{A}}$, we can now set up two master equations from which the two unknown coefficients can be obtained. The first equation is based on the evaluation of the $\hat{A}^2$ operator, which must yield the result from Eq. (12.69):

$$\begin{aligned} \langle nlm|\,\hat{A}^2\,|nlm\rangle &= \langle nlm|\left(\hat{A}_+\hat{A}_- + \hat{A}_z^2 - \hat{L}_z\right)|nlm\rangle \\ &= \left\langle \hat{A}_- nlm|\hat{A}_- nlm\right\rangle + \left\langle \hat{A}_z nlm|\hat{A}_z nlm\right\rangle - m \\ &= n^2 - 1 - l(l+1). \end{aligned} \quad (12.96)$$

Substituting the matrix elements then yields the first master equation:

$$\begin{aligned} &\left(\frac{c^+_{nlm}}{2l+1}\right)^2\left(\frac{l-m+2}{l+m+1}+1\right) + \left(\frac{c^-_{nlm}}{2l+1}\right)^2\left(\frac{l+m-1}{l-m}+1\right) \\ &\qquad = n^2 - 1 - l(l+1) + m. \end{aligned} \quad (12.97)$$

For the second equation, let us express the commutator of the raising and lowering operators in terms of the c coefficients. This commutator reads $\left[\hat{A}_+,\hat{A}_-\right] = 2\hat{L}_z$. Putting this result inside a matrix element yields

$$\begin{aligned} &\langle nlm|\left(\hat{A}_+\hat{A}_- - \hat{A}_-\hat{A}_+ - 2\hat{L}_z\right)|nlm\rangle \\ &\quad = \left\langle \hat{A}_- nlm|\hat{A}_- nlm\right\rangle - \left\langle \hat{A}_+ nlm|\hat{A}_+ nlm\right\rangle - 2m = 0. \end{aligned} \quad (12.98)$$

On substitution, we obtain

$$\begin{aligned} &\left(\frac{c^+_{nlm}}{2l+1}\right)^2\left(\frac{l-m+2}{l+m+1} - \frac{l+m+2}{l-m+1}\right) \\ &\quad + \left(\frac{c^-_{nlm}}{2l+1}\right)^2\left(\frac{l+m-1}{l-m} - \frac{l-m-1}{l+m}\right) - 2m = 0. \end{aligned} \quad (12.99)$$

This set of two equations yields the absolute values of the c coefficients. At this point, a phase choice must be made to obtain a fixed form for the raising and lowering operations. We have adopted here the phase choice made by Adams[12] and colleagues,

---

[12] B. G. Adams, J. Čížek, and J. Paldus. "Representation Theory of SO(4, 2) for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods." 1982.

Table 12.6 Action of the step operators on the $|nlm\rangle$ ket functions. Notice that $c_l = \frac{\sqrt{n^2-l^2}}{\sqrt{4l^2-1}}$.

$$\hat{L}^2\,|nlm\rangle = \hbar^2 l\,(l+1)\,|nlm\rangle$$

$$\hat{L}_z\,|nlm\rangle = \hbar m\,|nlm\rangle$$

$$\hat{L}_+\,|nlm\rangle = \hbar\sqrt{(l-m)\,(l+m+1)}\,|nl\,(m+1)\rangle$$

$$\hat{L}_-\,|nlm\rangle = \hbar\sqrt{(l+m)\,(l-m+1)}\,|nl\,(m-1)\rangle$$

$$\hat{Q}^2\,|nlm\rangle = \hbar^2 l\,(l+1)\,|nlm\rangle$$

$$\hat{Q}_z\,|nlm\rangle = \hbar n\,|nlm\rangle$$

$$\hat{Q}_+\,|nlm\rangle = \hbar\sqrt{(n-l)\,(n+l+1)}\,|(n+1)\,lm\rangle$$

$$\hat{Q}_-\,|nlm\rangle = \hbar\sqrt{(n+l)\,(n-l-1)}\,|(n-1)\,lm\rangle$$

$$\hat{A}^2\,|nlm\rangle = \hbar^2\left(n^2 - l\,(l+1) - 1\right)|nlm\rangle$$

$$\hat{A}_z\,|nlm\rangle = \hbar\sqrt{(l-m)\,(l+m)}c_l\,|n\,(l-1)\,m\rangle + \hbar\sqrt{(l-m+1)\,(l+m+1)}c_{l+1}\,|n\,(l+1)\,m\rangle$$

$$\hat{A}_+\,|nlm\rangle = \hbar\sqrt{(l-m)\,(l-m-1)}c_l\,|n\,(l-1)\,(m+1)\rangle - \hbar\sqrt{(l+m+2)\,(l+m+1)}c_{l+1}\,|n\,(l+1)\,(m+1)\rangle$$

$$\hat{A}_-\,|nlm\rangle = -\hbar\sqrt{(l+m)\,(l+m-1)}c_l\,|n\,(l-1)\,(m-1)\rangle + \hbar\sqrt{(l-m+2)\,(l-m+1)}c_{l+1}\,|n\,(l+1)\,(m-1)\rangle$$

yielding

$$\begin{aligned} \frac{c^{+}_{nlm}}{2l+1} &= \sqrt{(l+1-m)\,(l+1+m)}\sqrt{\frac{n^2-(l+1)^2}{4\,(l+1)^2-1}}; \\ \frac{c^{-}_{nlm}}{2l+1} &= \sqrt{(l-m)\,(l+m)}\sqrt{\frac{n^2-l^2}{4l^2-1}}. \end{aligned} \tag{12.100}$$

Note that the second expression in this equation is immediately obtained from the first by simply replacing $l+1$ with l. This reflects the unitary relationship between the $\hat{\Delta}^{+}_{l}$ and $\hat{\Delta}^{-}_{l+1}$ operators. Finally, Table 12.6 summarizes the action of all the step operators that allows us to change the hydrogenic quantum numbers.

13 The periodic table

The elements, if arranged according to their atomic weights, exhibit an apparent periodicity of properties.
–Dmitrii Ivanovich Mendeleev (1869)[1]

The *periodic system* represents a classification of the manifold of chemical elements. It was proposed by the Russian chemist Dmitrii Ivanovich Mendeleev (1834–1907) on February 17, 1869.[2] In addition to bringing order among the plethora of chemical and physical properties, the periodic table encouraged discussion in atomic physics and helped forging the early theories of quantum mechanics. As of 2017, Mendeleev's chart has been elevated to a polychromatic *icon* emblematic of the successes of modern science.

And yet, during the $\sim$150 years that have elapsed since Mendeleev's initial proposition, the overall structure of the periodic table has never been derived from

---

[1] D. I. Mendeleev. "On the Correlation Between the Properties of the Elements and Their Atomic Weights." *Zhurnal Russkogo Khimicheskogo Obshchestva* 1.2–3 (1869), pp. 35, 60–77.

[2] D. I. Mendeleev. *An Attempted System of the Elements Based on Their Atomic Weights and Chemical Analogies*. 1869. See: M. Kaji. "D. I. Mendeleev's Concept of Chemical Elements and the *Principles of Chemistry*." *Bulletin for the History of Chemistry* 27.1 (2002), pp. 4–16. D. I. Mendeleev. "Versuch eines Systems der Elemente nach ihren Atomgewichten und chemischen Funktionen." *Journal für praktische Chemie* 106 (1869), p. 251; D. I. Mendeleev. "Über die Beziehungen der Eigenschaften zu den Atomgewichten der Elemente." *Zeitschrift für Chemie* 5 (1869), pp. 405–406; D. I. Mendeleev. "Die periodischen Gesetzmässigkeit der chemischen Elemente." *Annalen der Chemie und Pharmacie* 8 (Suppl.) (1871), pp. 133–229.

first principles, nor has the origin of the periodic law been completely understood from a *quantum mechanical* perspective.

Chapter outline

This chapter provides a general overview of the current state of our understanding of the periodic law. The *chemical* periodicity phenomena that induced Mendeleev to construct his periodic table will be introduced in §13.1. Special emphasis will be laid, however, on the explanations provided by the latest *physical* and *mathematical* research. A brief review of the quantum mechanics of atomic systems appears in §13.2, and is followed in §13.3 by a discussion of the periodic system's overall structure from an atomic physics point of view.

Two crucially important features of the Mendeleev chart will be revealed in this way: (1) the *Madelung* $(n+l, n)$ *rule*, which is believed to rationalize the orbital filling order in many-electron systems and (2) the *doubling of the periods*, which emerges as a natural consequence of the Madelung rule. Both concepts are of such paramount importance to the periodic table that an entire section. §13.4, will be devoted to their study. A modified format of the periodic system with a distinctive *stepped* profile will be proposed on their basis, and will be referred to as the *left-step* or *Janet Periodic Table*.

One of the principal reasons for the current lack of understanding of the periodic law is the absence of an *ab initio* derivation of the Madelung rule. Many claims for a successful derivation have appeared in the scientific literature, but most have been dismissed. As a result, the $(n+l, n)$ rule has fallen into disrepute among many chemists and physicists alike. Its utility in describing the electronic structure of neutral atoms has been called into question, and its universal validity has been disputed. Indeed, most quantum mechanical interpretations of the $(n+l, n)$ rule are found lacking in many respects, as will be explained in §13.5.

Instead therefore, a *group theoretical* articulation of the $(n+l, n)$ rule will be proposed in this chapter. A concise and nontechnical description of this approach will be provided in §13.6. Both the philosophical and methodological implications of this approach will also be discussed. This will set the scene for a detailed treatment of the symmetries of the periodic table in Chapter 14.

13.1 CHEMICAL PERIODICITY

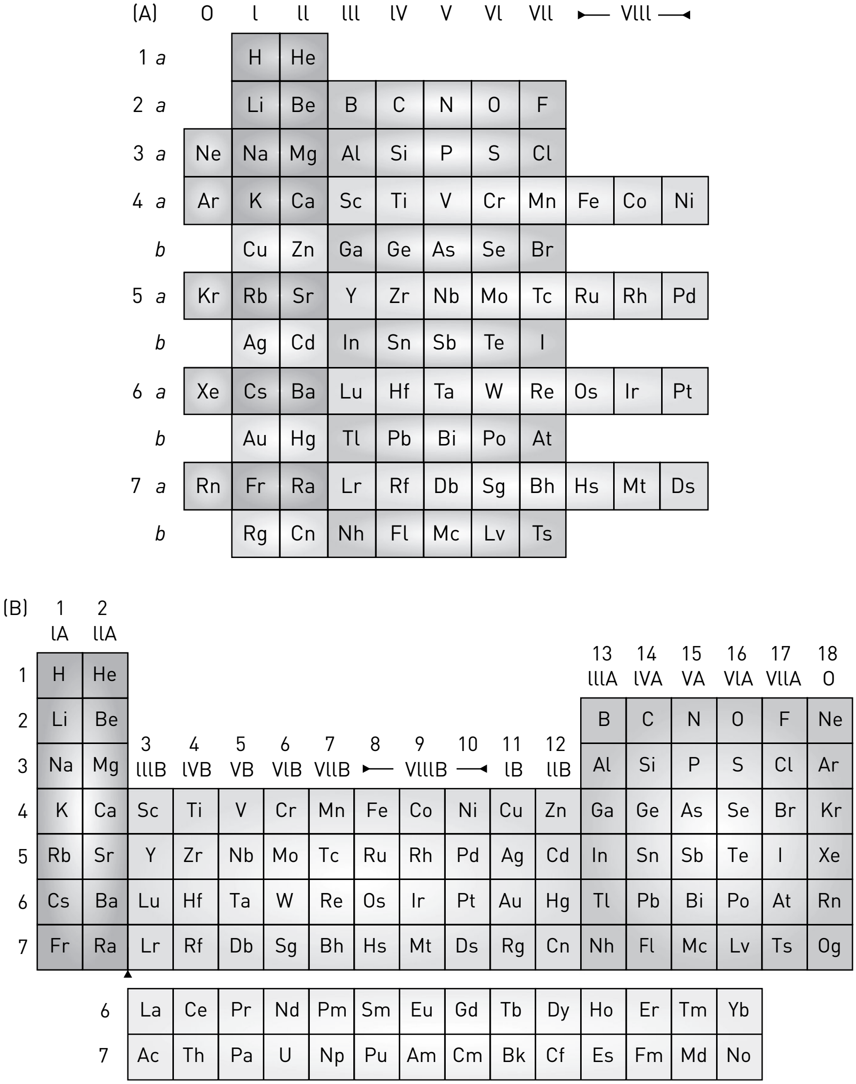
A two-step process is needed to build a periodic table from scratch. First, all the elements have to be ordered according to increasing atomic number Z. This *primary classification* results in a long horizontal sequence of elements, and has been called the *Mendeleev line*.[3] It will be noted that certain chemical and physical properties of the elements recur *periodically* (i.e., after regular intervals). These *periodicity phenomena* are best exhibited by plotting some chemical or physical property of the elements versus the ordinal (atomic) number Z. Typical examples include plots of the atomic radii, electronegativities, or first ionization energies of the atoms.

---

[3] H. A. Bent. *New Ideas in Chemistry from Fresh Energy for the Periodic Law*. Bloomington, IN: AuthorHouse, 2006.

In a second step, termed the *secondary classification*, the Mendeleev line is partitioned at certain well-defined loci, and the resulting sections (i.e., *periods*) are then placed underneath each other so that elements with similar properties fall into the same vertical columns, thus forming *natural groups* and representing the periodic law graphically (Figure 13.1). In this sense, the periodic system of the elements can be said to *embody* the periodic law.

Definition 13.1 (The periodic law): The physical and chemical properties of the chemical elements recur periodically as their atomic number increases. ▪

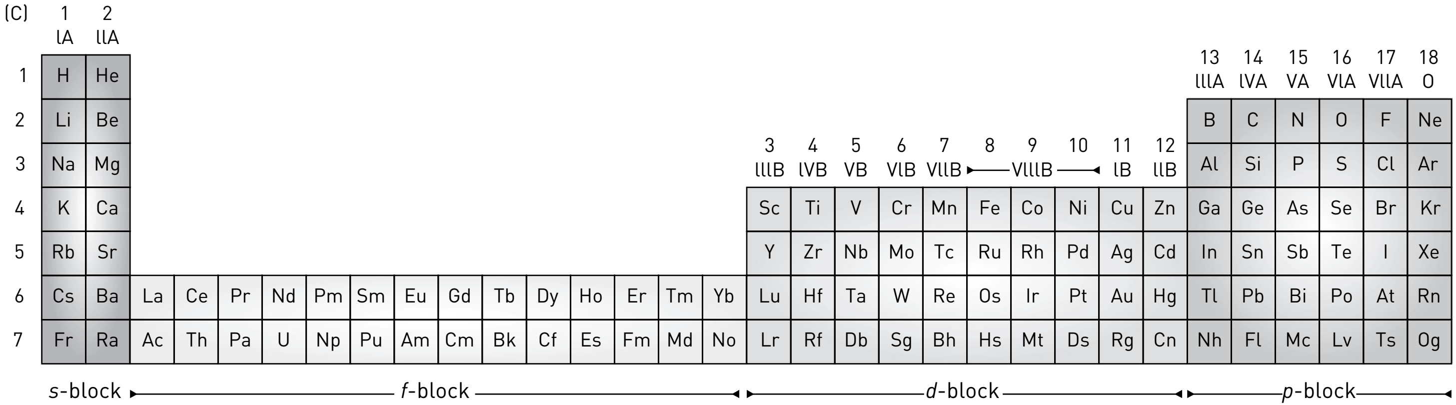


(A)	0	I	II	III	IV	V	VI	VII	►— VIII —◄		
1 *a*		H	He								
2 *a*		Li	Be	B	C	N	O	F			
3 *a*	Ne	Na	Mg	Al	Si	P	S	Cl			
4 *a*	Ar	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni
b		Cu	Zn	Ga	Ge	As	Se	Br			
5 *a*	Kr	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd
b		Ag	Cd	In	Sn	Sb	Te	I			
6 *a*	Xe	Cs	Ba	Lu	Hf	Ta	W	Re	Os	Ir	Pt
b		Au	Hg	Tl	Pb	Bi	Po	At			
7 *a*	Rn	Fr	Ra	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds
b		Rg	Cn	Nh	Fl	Mc	Lv	Ts			

(B)	1 IA	2 IIA	3 IIIB	4 IVB	5 VB	6 VIB	7 VIIB	8 ►—	9 VIIIB	10 —◄	11 IB	12 IIB	13 IIIA	14 IVA	15 VA	16 VIA	17 VIIA	18 O
1	H	He																
2	Li	Be											B	C	N	O	F	Ne
3	Na	Mg											Al	Si	P	S	Cl	Ar
4	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
5	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
6	Cs	Ba	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
7	Fr	Ra	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og

6	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb
7	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No

FIGURE 13.1



(C)	1 IA	2 IIA															3 IIIB	4 IVB	5 VB	6 VIB	7 VIIB	8	9 VIIIB	10	11 IB	12 IIB	13 IIIA	14 IVA	15 VA	16 VIA	17 VIIA	18 0
1	H	He																														
2	Li	Be																									B	C	N	O	F	Ne
3	Na	Mg																									Al	Si	P	S	Cl	Ar
4	K	Ca															Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
5	Rb	Sr															Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
6	Cs	Ba	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
7	Fr	Ra	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og
	s-block		*f*-block														*d*-block										*p*-block					

FIGURE 13.1 Short-form (A), medium-longform (B), and long-form (C) depiction of the periodic table. Notice that the periods in the short form table have been subdivided in *a* and *b* series (the lanthanides and actinides are not included).

Quasiperiodicity and period lengths

The observed periodicity is of a *quasiperiodic* nature. Systems exhibiting mathematical (instead of chemical) periodicity are typically characterized by a single period length.[4] Mendeleev's system, on the contrary, contains period lengths that vary and increase along the table. The standard periodic table (Figure 13.1B), for instance, consists of seven periods with the following *cardinalities* (i.e., number of elements within a period)[5]:

$$2-8-8-18-18-32-32. \tag{13.1}$$

The set in Eq. (13.1) represents a crucial feature of the periodic system. Notice also that all period lengths occur in pairs, except for the very first period of dimension 2. We will come back to this point in §13.4.

13.2 QUANTUM MECHANICS OF ATOMIC SYSTEMS

The construction of the periodic table by Mendeleev revealed the existence of an integer number regularity [Eq. (13.1)], which hinted at the compound nature of the atom. This was an open invitation to explore the deeper structure of atomic matter. It would, nonetheless, take several decades before the idea of the indivisible atom was abandoned and nuclei and electrons were identified as the atomic constituents.

With the advent of quantum mechanics at the beginning of the twentieth century, the aforementioned periodicity phenomena received a quantum mechanical underpinning. Historically, the first pioneering work in this direction was pursued by Niels Bohr (1885–1962). As early as 1913, Bohr used the *old quantum theory* to provide a theoretical explanation for the periodicity phenomena on the basis of the electronic configurations of atoms. Most of his theoretical findings have now entered mainstream textbook knowledge. However, on a deeper level, many characteristic features of the periodic system are still in need of explanation.

The goal of this section is to provide a brief summary of the (nonrelativistic) quantum mechanics of atomic systems. We will build up the complexity by starting with a consideration of one-electron (hydrogenic) systems in §13.2.1, which will set the scene for a treatment of *many-electron systems* in §13.2.2. This will provide the necessary background to tackle the periodic table in the next section, §13.3.

13.2.1 One-electron systems

Recall that the energy eigenvalues of hydrogenic systems (H, He^+, Li^{++}, and so on) are quantized according to the following energy formula:

$$E_n = -\frac{mZ^2 e^4}{8h^2 {\epsilon_0}^2 n^2} \quad \text{for} \quad n = 1, 2, 3, 4, \ldots, \tag{13.2}$$

where m is the reduced mass, Z is the atomic number ($Z = 1$ for hydrogen), e is the unit of electric charge, h is Planck's constant (the logo of quantum mechanics), and ϵ_0

---

[4] Think of sinusoidal functions, a planet tracing its orbit around the sun, the start of a new day every twenty-four hours, or the return of the winter season every new year.

[5] The completion of the seventh period dates from December 30, 2015, when the International Union of Pure and Applied Chemistry (IUPAC) announced in a Press Release the discovery and assignment of elements with atomic numbers 113, 115, 117, and 118.

is the vacuum permitivity. It is evident that this spectrum is *many-fold degenerate* in l and m_l, dependent only on the quantum number n (Figure 9.1). The total degeneracy (or *multiplicity*) for the principal quantum number n is equal to n^2, according to

$$\sum_{l=0}^{n-1}(2l+1) = (0+(n-1))\,n + \sum_{l=0}^{n-1} 1 = (n-1)\,n + n = n^2. \tag{13.3}$$

The energy rule in Eq. (13.2) gives rise to the following orbital sequence for the hydrogen atom (taking into account the spin degeneracy, which doubles the multiplicity to $2n^2$):

$$\underbrace{\overbrace{\{1s\}}^{n=1}}_{\dim=2} \ll \underbrace{\overbrace{\{2s=2p\}}^{n=2}}_{\dim=8} \ll \underbrace{\overbrace{\{3s=3p=3d\}}^{n=3}}_{\dim=18} \ll \underbrace{\overbrace{\{4s=4p=4d=4f\}}^{n=4}}_{\dim=32} \ll \ldots. \tag{13.4}$$

We can summarize this sequence of levels of increasing energy by the simple *Fock (n) rule.*

Definition 13.2 (The Fock (n) rule): With increasing nuclear charge Z, the nl-orbitals are filled in order of increasing n. ▪

13.2.2 Many-electron systems

The quantum mechanical description of a multielectron atom (with atomic number Z) is compounded by the inception of nuclear *shielding effects* and *interelectronic interactions* which dramatically change the hydrogenic order (Eq. (13.4)) of nl orbitals.[6] Moreover, these factors render the exact analytic solution of the Schrödinger equation

$$\hat{\mathscr{H}}\Psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_Z) = E\Psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_Z), \tag{13.5}$$

with

$$\hat{\mathscr{H}} = -\sum_{i=1}^{Z}\frac{\hbar^2}{2m}\nabla_i^2 - \sum_{i=1}^{Z}\frac{Ze^2}{r_i} + \sum_{i=1}^{Z-1}\sum_{j>i}^{Z}\frac{e^2}{r_{ij}}, \tag{13.6}$$

impossible. Hence, the theoretical chemist is forced to use approximate numerical methods. The first term in Eq. (13.6) contains the kinetic energies for the Z electrons, the second sum is the potential energy for the attractions of each electron toward the nucleus of charge Ze, and the last term gives the potential energy of the electronic repulsions. Notice that an infinitely heavy point nucleus is assumed.

The orbital approximation

Within the *hydrogenic limit* (or *single-particle approach*), the Z electrons are assumed to behave independently of one another by suppressing the interelectronic repulsion

---

[6] For a *hydrogenic system* (H, He^+, Li^{++}, and so on), the energy is independent of l so that levels with larger values of n invariably have higher energies. For *many-electron systems*, on the other hand, the energy is dependent on both n and l, and a level with $n=4$ and $l=0$ (i.e., $4s$-orbital), for example, can lie below the level with $n=3$ and $l=2$ (i.e., $3d$-orbital). See §13.3.

term in Eq. (13.6):

$$\hat{\mathscr{H}} = \sum_{i=1}^{Z}\left(-\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Ze^2}{r_i}\right) = \sum_{i=1}^{Z}\hat{\mathscr{H}}_i. \tag{13.7}$$

The total wave function Ψ can then be treated as a determinental product of Z one-electron wave functions ψ_i (*spin orbitals*).[7]

The self-consistent field approximation

The *self-consistent field* approximation assumes that each electron moves in the average, self-consistent field created by the other electrons and the atomic nucleus. As a result, it is actually subject to an *effective one-electron potential* $V_{\text{eff}}(r_i)$. On account of the spherical symmetry of this potential, this approach is also known as the *central field approximation*. To each orbital ψ_i then corresponds a particular energy value, denoted ε_{nl}, according to the analytically solvable time-independent Schrödinger equation for one-electron systems[8]:

$$\hat{\mathscr{H}}_i\psi_i(\mathbf{r}_i) = \left[-\frac{\hbar^2}{2m}\nabla_i^2 + V_{\text{eff}}(r_i)\right]\psi_i(\mathbf{r}_i) = \varepsilon_{nl}\psi_i(\mathbf{r}_i) \tag{13.8}$$

Note that a *screening function* $\zeta(r_i)$ can be defined that specifically accounts for the screening of the nuclear Coulomb potential $-Ze^2/r_i$ by writing V_{eff} in the form

$$V_{\text{eff}}(r_i) = -\zeta(r_i)\frac{Ze^2}{r_i}. \tag{13.9}$$

Following Bohr's *Aufbauprinzip* (or *building-up principle*), an atom of atomic number Z is considered to be the end result of a process of adding Z electrons in succession to the bare atomic nucleus. During this building-up process, the electrons tend to occupy the one-electron levels ψ_i of lowest energy ε_{nl} first. Because the number of electrons per orbital is limited by Pauli's *exclusion principle*, more and more spin orbitals are progressively occupied with increasing atomic number Z.[9] The order in which these orbitals are sequentially occupied is known as Bohr's *Aufbau scheme*, and the final orbital occupation is referred to as the atom's electronic *ground state configuration*.

13.3 QUANTUM MECHANICS OF THE PERIODIC SYSTEM

Having reviewed the quantum mechanics of atomic systems briefly in §13.2, all ingredients are in place for a quantum mechanical description of the periodic system. This description is based on the following three principles:

---

[7] The total wave function Ψ is also called the *Hartree wave function*.

[8] The atomic energy levels ε_{nl} are also referred to as the *spectral terms* of the atom.

[9] As a reminder, the Pauli exclusion principle follows from the *indistinguishability* of the electrons (or fermions, in general). That is, the act of exchanging two electrons should remain hidden from direct observation. A more accurate statement of the Pauli principle therefore enounces that the wave function Ψ should be *antisymmetrized* under the permutation of electrons. The electrons consequently lose their *individuality*, and the quantum numbers n, l, m_l, and m_s no longer "belong" to any individual electron, but label the one-electron orbitals ψ_i instead. Ostrovsky thus stressed that "using atomic configurations *does not* mean that an orbital, or set of quantum numbers are ascribed to any particular electron." See V. N. Ostrovsky. "What and How Physics Contributes to Understanding the Periodic Law." *Foundations of Chemistry* 3.2 (2001), pp. 145–182.

1. *The quantum numbers n and l*: Many-electron atoms are characterized by a set of principal and orbital quantum numbers $\{n, l\}$ that label occupied one-electron wave functions (orbitals).
2. *The Pauli exclusion principle*: No two electrons can share the same set of four quantum numbers n, l, m_l, and m_s. Each orbital can accommodate at most two electrons, which differ in their spin projection $m_s = \pm 1/2$: one with spin up and the other with spin down.
3. *The Aufbau principle*: One-electron wave functions are filled sequentially with electrons in order of increasing energy.

Based on these three principles, the *ground state configurations* of the elements can be determined, and a periodic table can be constructed in which elements with the same electronic valence shell configuration are grouped together into the same column. The periodic system thus *reflects* the order of filling of the electron nl (sub)shells.

The ground state configuration of each element is dependent on the nature of the Aufbau scheme, the outline of which is given by the *energetic distribution* of the one-electron levels within a neutral atom. This energetic order is determined by the form of the atomic potential V_{eff} in Eq. (13.8). A variety of periodic tables with radically different periodicity patterns (i.e., set of period lengths) can thus be obtained, depending on the exact form of the potential V_{eff}.[10]

The search for a generalized effective potential V_{eff} that rationalizes the structure of the standard periodic table has posed severe challenges to quantum chemists. To arrive at an insightful explanation for the oddly shaped periodic chart, it will be important to keep the subject matter simple. In the following sections, three effective one-electron potentials V_{eff} of increasing complexity will be reviewed, and we will ascertain whether the corresponding Aufbau schemes (or *energy ordering rules*) are capable of yielding the correct sequence of energy levels to account for the structure of the periodic system.

13.3.1 The Fock (*n*) rule

Many chemists intuitively assume $V_{\text{eff}}(\mathbf{r})$ to be described by the pure hydrogenic Coulomb potential $-Ze/r$. This yields the Aufbau scheme in Eq. (13.4) as described by the Fock (n) rule (Definition 13.2).

Based on this scheme and the three principles, the electron in atomic hydrogen (H, $Z = 1$) goes to the 1s-orbital with $n = 1$, $l = 0$, $m_l = 0$, and $m_s = +1/2$. Its *ground state configuration* is therefore $1s^1$. Moving on to helium (He, $Z = 2$), a second electron with opposite spin (i.e., $m_s = -1/2$) is added to the 1s-orbital, yielding the electron configuration $1s^2$. This exhausts all the possibilities for the $n = 1$ state and results in a *closed shell* (a shell with its full complement of electrons), denoted [He] when it occurs as a core in other atoms.

Next comes lithium (Li, $Z = 3$), which has three electrons. After the first two electrons huddle in the $n = 1$ state, the third electron must go into the $n = 2$ state. But there, a problem arises! According to the Aufbau scheme (Eq. (13.4)), both the 2s-orbital and the three 2p-orbitals have the same energy. So, the question arises whether the third electron should occupy a 2s-orbital or one of the three 2p-orbitals?!

---

[10] ibid., p. 154, and V. N. Ostrovsky. "Physical Explanation of the Periodic Table." *Annals of the New York Academy of Sciences* 988 (2003), pp. 182–192.

13.3.2 The hydrogenic (n, l) rule

The Fock (n) rule thus fails to account for the ground state electronic configurations of the elements. It does so because it completely ignores the presence of other electrons that typically *screen* the nuclear charge. In a first, crude approximation, we could therefore assume the Coulomb potential to be slightly *distorted*; that is, we could postulate that V_{eff} corresponds to a *weakly perturbed* Coulomb potential. Let us see how far this assumption can lead us in the next few sections.

Screening effects

The energy with which an electron is held in a *hydrogenic* atom is, we repeat, proportional to Z^2/n^2:

$$E_n = -\frac{mZ^2e^4}{8h^2\varepsilon_0{}^2n^2} \quad \text{for} \quad n = 1, 2, 3, 4, \ldots. \tag{13.10}$$

This formula can be used to calculate the nth *ionization energy* $I_n(\mathrm{X})$ of a many-electron atom X in the gaseous phase. For the ground state configuration of the hydrogen atom $(Z = 1)$, for instance, the electron is located in the $1s$-orbital $(n = 1)$, with an energy of $E_1 = -13.6$ eV, thus yielding an ionization energy of $I_1(\mathrm{H}) = 13.6$ eV. This value is remarkably close to the experimentally obtained ionization energy of hydrogen $I_1(\mathrm{H}) = 13.59844$ eV (see Table 13.1).

Similar calculations can be performed on the other elements. Helium $(Z = 2)$ and lithium $(Z = 3)$, for example, have their outermost valence electron located in the $1s$- and $2s$-orbital, respectively, which yields the following ionization energies:

$$I_1(\mathrm{He}) = \frac{4}{1}\, I_1(\mathrm{H}) = 54.4 \text{ eV}; \tag{13.11}$$

$$I_1(\mathrm{Li}) = \frac{9}{4}\, I_1(\mathrm{H}) = 30.6 \text{ eV}. \tag{13.12}$$

Table 13.1 Calculated (calc.) and observed (obs.) ionization energies for the first five elements: hydrogen $(Z = 1)$, helium $(Z = 2)$, lithium $(Z = 3)$, beryllium $(Z = 4)$, and boron $(Z = 5)$. The discrepancy between the observed and calculated values is due to a *screening* of the nuclear charge by the other electrons. The *effective nuclear charge* Z_{eff} and *shielding constant* σ are listed for the outermost valence electron.

		I_1	I_2	I_3	I_4	I_5	Z_{eff}	σ
H	calc.	13.6 eV	—	—	—	—	1.00	0.00
	obs.	13.6 eV	—	—	—	—	1.00	0.00
He	calc.	54.4 eV	54.4 eV	—	—	—	2.00	0.00
	obs.	24.6 eV	54.4 eV	—	—	—	1.34	0.66
Li	calc.	30.6 eV	122.4 eV	122.4 eV	—	—	3.00	0.00
	obs.	5.39 eV	75.6 eV	122.5 eV	—	—	1.26	1.74
Be	calc.	54.4 eV	54.4 eV	217.6 eV	217.6 eV	—	4.00	0.00
	obs.	9.32 eV	18.2 eV	153.9 eV	217.7 eV	—	1.65	2.35
B	calc.	85.0 eV	85.0 eV	85.0 eV	340.0 eV	340.0 eV	5.00	0.00
	obs.	8.30 eV	25.2 eV	37.9 eV	259.4 eV	340.2 eV	1.56	3.44

Admittedly, such calculations only yield approximate values for I_n. Table 13.1 illustrates that the experimentally observed ionization energies are always less than the theoretically deduced values.

This discrepancy can be explained by the fact that in many-electron systems, the presence of the other electrons causes the atomic nucleus to be *screened*. Each electron therefore experiences an *effective nuclear charge*,[11] denoted Z_{eff}, which is smaller than the *actual nuclear charge* Z by a factor σ according to

$$Z_{\text{eff}} = Z - \sigma \leq Z, \tag{13.13}$$

where σ is the *shielding* or *screening constant* (see Table 13.1). The nuclear charge, as felt by the valence electrons of He, Li, and Be, for instance, gets reduced from 2, 3, and 4 to 1.34, 1.26, and 1.65, respectively. From this point of view, the shielding constant stands for the average number of electrons between the nucleus and the electron in question.[12]

The orbital penetration effect

Let us briefly consider the case of the lithium atom ($Z = 3$) with which we ended the last section, which has a groundstate configuration that is known to be $1s^2 2s^1$. Figure 13.2 exhibits the *radial probability distribution* for various hydrogenic atomic orbitals. It follows from this graph that the electron density of the 2*s* valence electron is situated predominantly outside the 1*s*-orbital, which suggests that the +3 nuclear charge of lithium is screened by two 1*s*-electrons, yielding a Z_{eff} of +1.[13] The reason why $Z_{\text{eff}}(\text{Li}) = 1.26$ rather than 1.00 (cf. Table 13.1) is because of another (smaller) bump in the radial probability plot of the 2*s*-orbital, the peak of which coincides with the maximum of the 1*s*-curve (see Figure 13.2). This implies that the 2*s* electron slightly *penetrates* the 1*s*-orbital, causing its effective nuclear charge to increase a little.

This *penetrating power P* of an orbital is dependent on the orbital quantum number *l*—a phenomenon referred to as the *orbital penetration effect*. The penetrating power is greatest for an *s*-orbital ($l = 0$) and decreases with increasing *l*: $P_s > P_p > P_d > P_f$. As a result of this variation in penetrating power, electrons with different *l*-values also experience different shieldings. An electron in an *s*-orbital, for example, is screened less effectively by the electrons in the other orbitals and it experiences a larger effective nuclear charge because of its highly penetrating power. This also means, conversely, that an *s* electron has greater *shielding power* than electrons in a *p*-, *d*-, or *f*-orbital of the same *n*-shell. As a general rule, the shielding constant σ increases with increasing l ($\sigma_s < \sigma_p < \sigma_d < \sigma_f$), and Z_{eff} decreases with increasing l according to Eq. (13.13).

---

[11] The effective nuclear charge is also called the *core charge* or *kernel charge*; the *kernel* being defined as the ensemble of the nucleus with the shielding electrons.

[12] Values of the shielding constant can be determined via the semiempirical *Slater rules* or via quantum mechanical calculations. See J. C. Slater. "Atomic Shielding Constants." *Physical Review* 36.1 (1930), pp. 57–64, and E. Clementi and D. L. Raimondi. "Atomic Screening Constants from SCF Functions." *The Journal of Chemical Physics* 38.11 (1963), pp. 2686–2689, as well as E. Clementi, D. L. Raimondi, and W. P. Reinhardt. "Atomic Screening Constants from SCF Functions: II. Atoms with 37 to 86 Electrons." *The Journal of Chemical Physics* 47.4 (1967), pp. 1300–1307.

[13] In electrostatics, Gauss's law shows that for a point outside a charged sphere, the entire electrical charge can be treated as a point charge situated at the origin of the sphere.

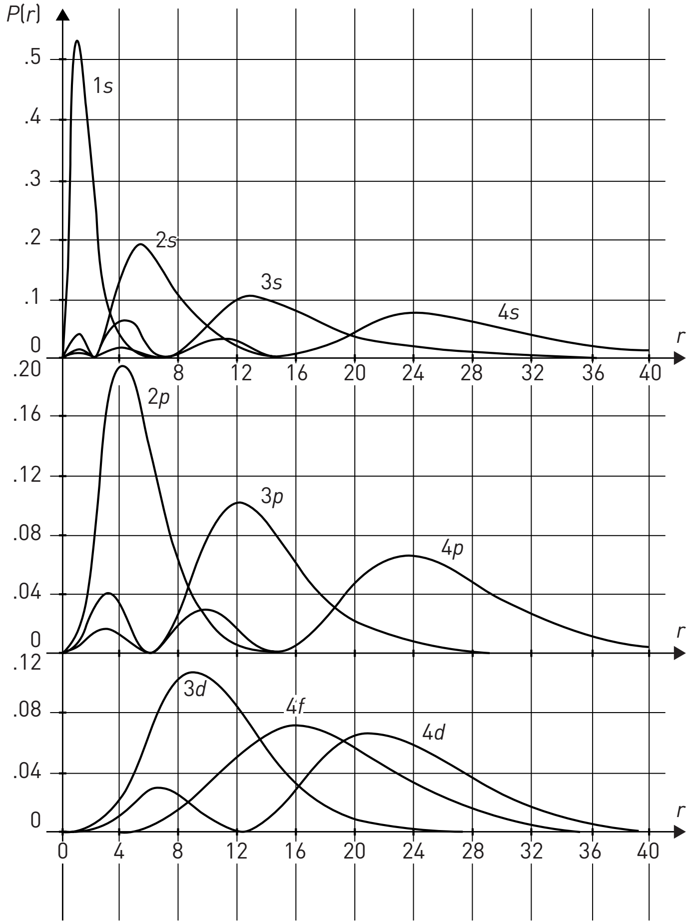


FIGURE 13.2 Radial probability distribution for various hydrogenic atomic orbitals. The radial probability $4\pi r^2 R_{n,l}^2$ has been plotted as a function of the radial distance r from the nucleus in atomic units.

The hydrogenic (*n, l*) rule

As a consequence of the previously mentioned screening effects, the degeneracy in the orbital quantum number l is lifted for orbitals in multielectron atoms and, as a general rule, $\varepsilon_s < \varepsilon_p < \varepsilon_d < \varepsilon_f$ within an n-shell. This leads to an important modification of the (Fock) orbital sequence in Eq. (13.4):

$$\underbrace{\overbrace{\{1s\}}^{n=1}}_{\text{dim}=2} \ll \underbrace{\overbrace{\{2s < 2p\}}^{n=2}}_{\text{dim}=8} \ll \underbrace{\overbrace{\{3s < 3p < 3d\}}^{n=3}}_{\text{dim}=18} \ll \underbrace{\overbrace{\{4s < 4p < 4d < 4f\}}^{n=4}}_{\text{dim}=32} \ll \ldots, \tag{13.14}$$

which can be summarized by the *hydrogenic* (n, l) *rule.*

Definition 13.3 (The hydrogenic (n, l) rule): With increasing nuclear charge Z, the orbitals are filled in order of increasing n; for fixed n, the orbitals are filled according to increasing l. ■

The first part of this rule corresponds to the Fock (n) rule for a pure, unscreened Coulomb field, whereas the second part accounts for the screening effects that lift the degeneracy over l. The orbitals in Eq. (13.14) are grouped according to the same value of n. Taking into account the possible values for the magnetic and spin magnetic quantum numbers m_l and m_s, this leads to the following series of dimensionalities for the above sequence:

$$2-8-18-32-50-\ldots, \tag{13.15}$$

as summarized by the Stoner formula $2n^2$.

According to the Aufbau scheme in Eq. (13.14), the 2s-orbital is lying lower in energy than the 2p-orbital, which accounts for the [He] $2s^1$ configuration of lithium (Li, $Z=3$). As a matter of fact, the (n,l) rule correctly describes the electronic configurations of the first eighteen elements (H → Ar). However, it radically fails to account however for the ground state configurations of the elements with $Z \geq 19$. For example, based on the energy sequence in Eq. (13.14), we would expect the electronic configuration of potassium (K, $Z=19$) to be [Ne] $3s^2 3p^6 3d^1$, whereas spectroscopic research has revealed its configuration to be [Ne] $3s^2 3p^6 4s^1$! Clearly, then, the (n,l) rule "generates an incorrect periodicity pattern [for neutral atoms]" and another filling rule will have to be invoked.[14]

13.3.3 The Madelung $(n+l,n)$ rule

The reason behind the total breakdown of the (n,l) rule after $Z=18$ lies in the erroneous assumption that with the potential deepening at intermediate and small r, the corresponding lowering of the nl energy levels remains small enough so as not to affect the major n-grouping.[15] That is, although the typical order of energies $\varepsilon_{ns} < \varepsilon_{np} < \varepsilon_{nd} < \varepsilon_{nf}$ *within* an n-shell can be rationalized by assuming V_{eff} to be a *weakly* perturbed Coulomb potential, this simple (qualitative) picture falls short in yielding an explanation for the fact that $\varepsilon_{4s} < \varepsilon_{3d}$, for instance.

In reality, the effective one-electron potential V_{eff} for neutral atoms is seen to deviate so *strongly* from the pure Coulomb potential $-Ze/r$, that it would be better to describe it by a new (*non*-Coulomb) potential altogether. As a result, a dramatic rearrangement of the nl levels is effected with substantial *overlap* between levels of different n (thus explaining $\varepsilon_{4s} < \varepsilon_{3d}$, for instance). Despite this fact, "a new type of regularity emerges in the form of the *Madelung rule* [emphasis added]."[16]

Definition 13.4 (The Madelung $(n+l,n)$ rule): With increasing nuclear charge Z, one-electron orbitals are filled according to increasing $N=n+l$, being the sum of the principal quantum number n and the orbital quantum number l. For fixed N, the orbitals are filled in order of increasing n. ▪

The [Ne] $3s^2 3p^6 4s^1$ configuration of potassium can now be explained on the basis of the lower $n+l$ value for the 4s-orbital ($n=4$, $l=0 \rightarrow n+l=4$) as compared to the 3d-orbital ($n=3$, $l=2 \rightarrow n+l=5$). A full application of the $(n+l,n)$ rule gives rise

---

[14] Ostrovsky, "What and How Physics Contributes to Understanding the Periodic Law," p. 186.

[15] V. N. Ostrovsky. "The Periodic Table and Quantum Physics." *The Periodic Table: into the 21st Century*. Eds. D. H. Rouvray and R. B. King. Baldock, UK: Research Studies Press, 2004, pp. 331–370.

[16] Ostrovsky, "What and How Physics Contributes to Understanding the Periodic Law," p. 156.

Table 13.2 Application of the empirical $(n+l,n)$ rule (also known as the *Madelung rule*) according to which the orbitals in neutral atoms are filled in order of increasing $n+l$, and n for fixed $n+l$. The exact filling sequence is obtained by reading the quantum map from left to right and top to bottom.

$n+l$	$l=3$	$l=2$	$l=1$	$l=0$	N_{n+l}^{max}	$Z_i \rightarrow Z_f$	$X_i \rightarrow X_f$
1	—	—	—	$1s^2$	2	$1 \rightarrow 2$	H → He
2	—	—	—	$2s^2$	2	$3 \rightarrow 4$	Li → Be
3	—	—	$2p^6$	$3s^2$	8	$5 \rightarrow 12$	B → Mg
4	—	—	$3p^6$	$4s^2$	8	$13 \rightarrow 20$	Al → Ca
5	—	$3d^{10}$	$4p^6$	$5s^2$	18	$21 \rightarrow 38$	Sc → Sr
6	—	$4d^{10}$	$5p^6$	$6s^2$	18	$39 \rightarrow 56$	Y → Ba
7	$4f^{14}$	$5d^{10}$	$6p^6$	$7s^2$	32	$57 \rightarrow 88$	La → Ra
8	$5f^{14}$	$6d^{10}$	$7p^6$	$8s^2$	32	$89 \rightarrow 120$	Ac → 120

to the following orbital sequence (see also the data shown in Table 13.2):

$$\begin{gathered}
\overbrace{\underbrace{\{1s\}}_{\text{dim}=2}}^{n+l=1} \ll \overbrace{\underbrace{\{2s\}}_{\text{dim}=2}}^{n+l=2} \ll \overbrace{\underbrace{\{2p < 3s\}}_{\text{dim}=8}}^{n+l=3} \ll \overbrace{\underbrace{\{3p < 4s\}}_{\text{dim}=8}}^{n+l=4} \ll \overbrace{\underbrace{\{3d < 4p < 5s\}}_{\text{dim}=18}}^{n+l=5} \ll \\
\overbrace{\underbrace{\{4d < 5p < 6s\}}_{\text{dim}=18}}^{n+l=6} \ll \overbrace{\underbrace{\{4f < 5d < 6p < 7s\}}_{\text{dim}=32}}^{n+l=7} \ll \overbrace{\underbrace{\{5f < 6d < 7p < 8s\}}_{\text{dim}=32}}^{n+l=8} \ll \ldots,
\end{gathered} \tag{13.16}$$

with grouping according to constant $N = n + l$. This corresponds to the following series of repeated "double squares"[17]:

$$2 - 2 - 8 - 8 - 18 - 18 - 32 - 32 - \ldots . \tag{13.17}$$

Interestingly, when compared with Eq. (13.15), the hydrogenic dimensions are seen to appear exactly *twice* in the Madelung sequence—a phenomenon known as the *period doubling*. Both the Madelung rule and the period doubling are of such paramount importance to the periodic system that the next section will be devoted to their detailed description.

13.4 THE MADELUNG $(n+l,n)$ RULE

The $(n+l,n)$ rule predicts the onset of all atomic subshell occupations in the Aufbau sequence of the periodic system. That is, it effectively foretells the beginning of the transition metals, and lanthanide and actinide series. It rationalizes, in other words, the *overall structure* of the periodic system by providing us with a recipe for laying out the different s-, p-, d-, and f-blocks in the table. As a welcome extra, the Madelung rule also discloses the *period doubling* as an essential feature of the periodic system (§13.4.2). In this sense, the power of the $(n+l,n)$ rule lies in its *architectural description* of the Mendeleev chart.

---

[17] Per-Olov Löwdin. "Some Comments on the Periodic System of Elements." *International Journal of Quantum Chemistry* S3A (1969), pp. 331–334.

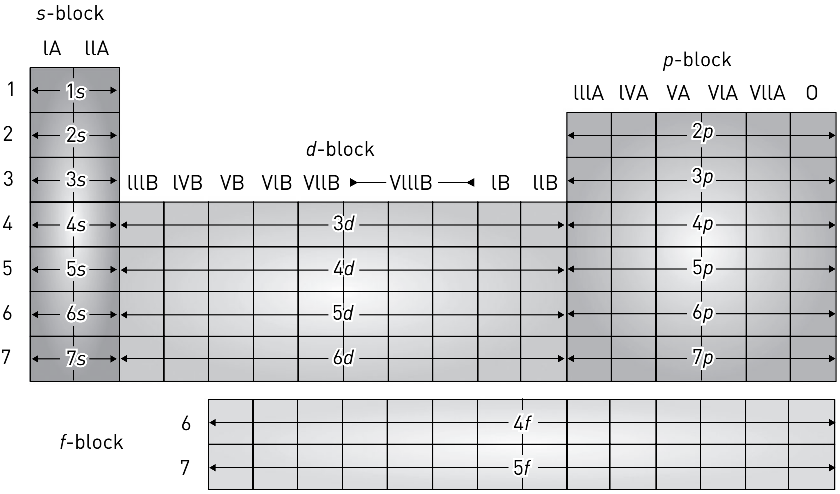


FIGURE 13.3 The overall stucture of the periodic table of chemical elements.

The previous section revealed the existence of a particular Madelung order (Eq. (13.16)) among the *nl* orbitals of many-electron systems. The Madelung $(n+l, n)$ rule thus holds the promise of yielding a deeper explanation for the periodicity phenomena described in §13.1. It is indeed sufficient to glance at the periodic table in Figure 13.3 to see that the Madelung $(n+l, n)$ rule faultlessly accounts for its *overall structure*. As Samuel A. Goudsmit and Paul I. Richards observed, the $(n+l, n)$ rule "is remarkably well obeyed throughout the periodic table."[18] It correctly accounts for the first occurrence of the atomic subshell occupations in the Aufbau sequence of the periodic system, predicting the inset of the transition metal block (3*d*-block) after the 4*s*-block, as well as the start of the lanthanide and actinide series (*f*-block elements) after the 6*s*- and 7*s*-elements, respectively.

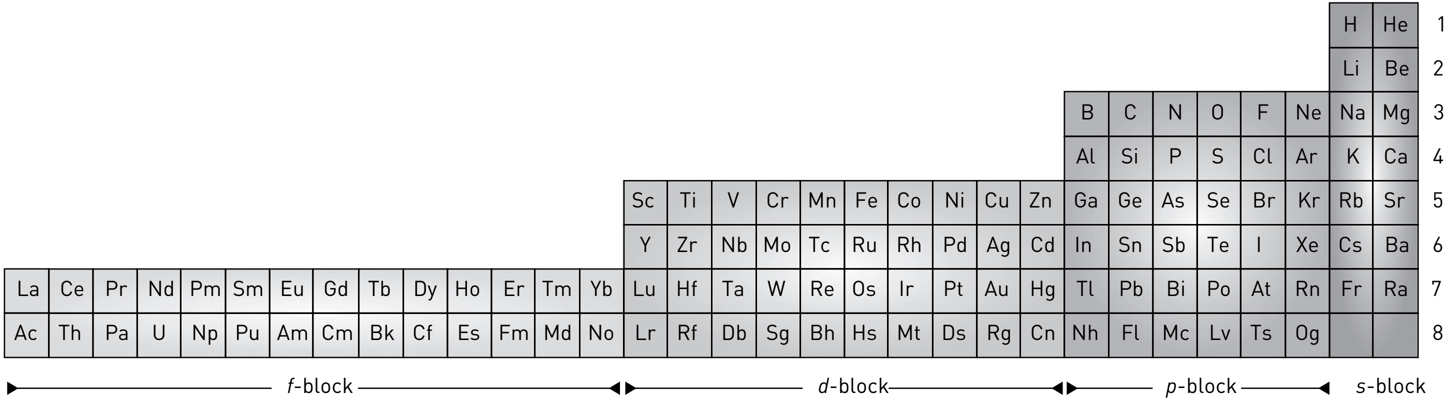
13.4.1 The left-step periodic table

The *quantum map* in Table 13.2 can be used as a framework for a new representation of the periodic system, as depicted in Figure 13.4. This form, known as the *eight-period* or *left-step periodic table* (LSPT), was first devised by the engineer and amateur biologist Charles Janet (1849–1932) in 1929 (Figure 13.5).[19] Janet had originally designed

---

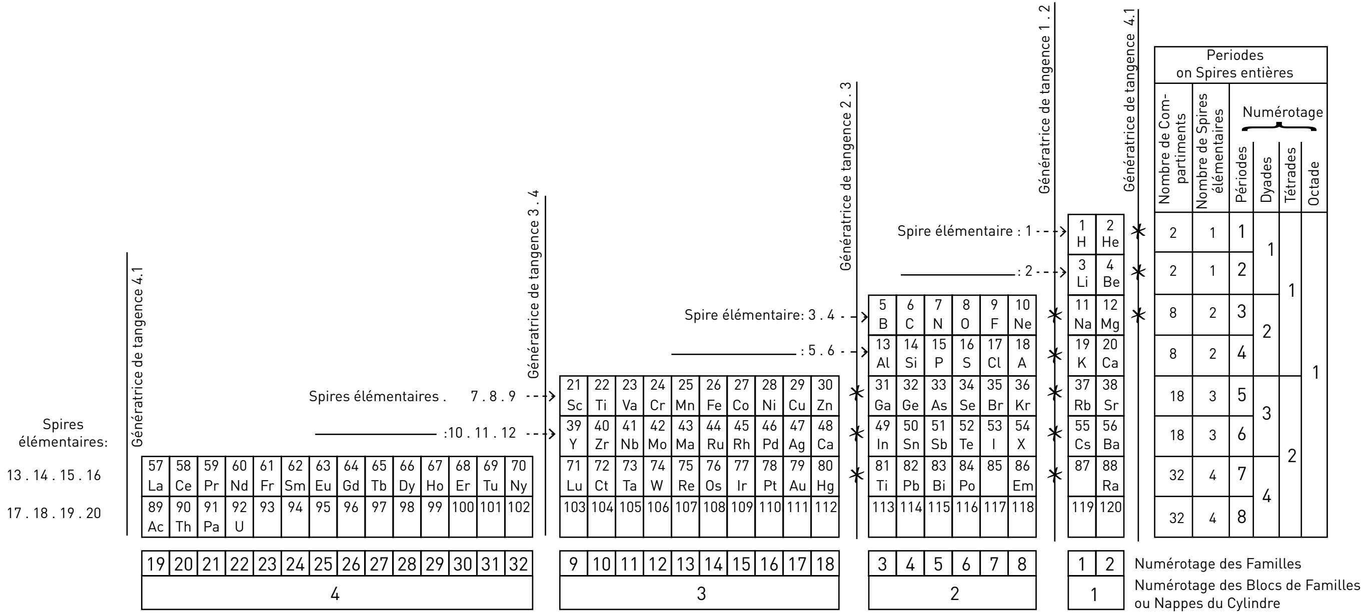
[18] S. A. Goudsmit and P. I. Richards. "The Order of Electron Shells in Ionized Atoms." *Proceedings of the National Academy of Sciences of the United States of America* 51 (1964), p. 664.

[19] C. Janet. Considérations sur la Structure du Noyau de l'Atome. Beauvais: Imprimerie Départementale de l'Oise, 1929. A preliminary helical version of the LSPT was published by Janet in 1927. See C. Janet. *La Structure du Noyau de l'Atome Considérée dans la Classification Périodique des Eléments Chimiques*. Beauvais: Imprimerie Départementale de l'Oise, 1927. This also explains Janet's preferred use of the term *hélicoïdale* (helicoid) for the LSPT, notwithstanding his later adoption of the adjective *scalariforme* (stepped). Various other geometric forms of the LSPT were designed in subsequent years. See, for example, C. Janet. *Essais de Classification Hélicoïdale des Èléments Chimiques*. Beauvais: Imprimerie Départementale de l'Oise, 1928, and C. Janet. *La Classification Hélicoïdale des Èléments Chimiques*. Beauvais: Imprimerie Départementale de l'Oise, 1928. For a detailed account of Janet's periodic systems, see P. J. Stewart. "Charles Janet: Unrecognized Genius of the Periodic System." *Foundations of Chemistry* 12 (2010), pp. 5–15. A biography of Janet



																														H	He	1
																														Li	Be	2
																								B	C	N	O	F	Ne	Na	Mg	3
																								Al	Si	P	S	Cl	Ar	K	Ca	4
														Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr	Rb	Sr	5
														Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe	Cs	Ba	6
La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn	Fr	Ra	7
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og			8
f-block														*d*-block										*p*-block						*s*-block		

FIGURE 13.4 Left-step periodic table. In this representation, the *f*-block consists of fourteen groups of *f*-elements with lanthanum (La) and actinium (Ac) as the first representatives of each row, and ytterbium (Yb) and nobelium (No) as the last ones. Lutetium (Lu) and lawrencium (Lr) are accommodated as *d*-block elements in the periodic table, below scandium (Sc) and yttrium (Y). The left-step periodic table is in perfect agreement with the *Madelung rule*.



Spires élémentaires																																	Periodes on Spires entières					
																																	Nombre de Com-partiments	Nombre de Spires élémentaires	Numérotage			
																																			Périodes	Dyades	Tétrades	Octade
1																															1 H	2 He	2	1	1	1	1	1
2																															3 Li	4 Be	2	1	2			
3 . 4																									5 B	6 C	7 N	8 O	9 F	10 Ne	11 Na	12 Mg	8	2	3	2		
5 . 6																									13 Al	14 Si	15 P	16 S	17 Cl	18 A	19 K	20 Ca	8	2	4			
7 . 8 . 9															21 Sc	22 Ti	23 Va	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr	37 Rb	38 Sr	18	3	5	3	2	
10 . 11 . 12															39 Y	40 Zr	41 Nb	42 Mo	43 Ma	44 Ru	45 Rh	46 Pd	47 Ag	48 Ca	49 In	50 Sn	51 Sb	52 Te	53 I	54 X	55 Cs	56 Ba	18	3	6			
13 . 14 . 15 . 16	57 La	58 Ce	59 Pr	60 Nd	61 Fr	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tu	70 Ny	71 Lu	72 Ct	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Ti	82 Pb	83 Bi	84 Po	85	86 Em	87	88 Ra	32	4	7	4		
17 . 18 . 19 . 20	89 Ac	90 Th	91 Pa	92 U	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	32	4	8			
Numérotage des Familles	19	20	21	22	23	24	25	26	27	28	29	30	31	32	9	10	11	12	13	14	15	16	17	18	3	4	5	6	7	8	1	2						
Numérotage des Blocs de Familles ou Nappes du Cylindre	4														3										2						1							

FIGURE 13.5 Charles Janet's left-step periodic table (1928). [From: C. Janet. *La Classification Hélicoidale des Eléments Chimiques.* 1928, Fig. 5. Reprinted with permission of the *Société Académique de l'Oise*, Beauvais, France. Courtesy of Loïc Casson.]

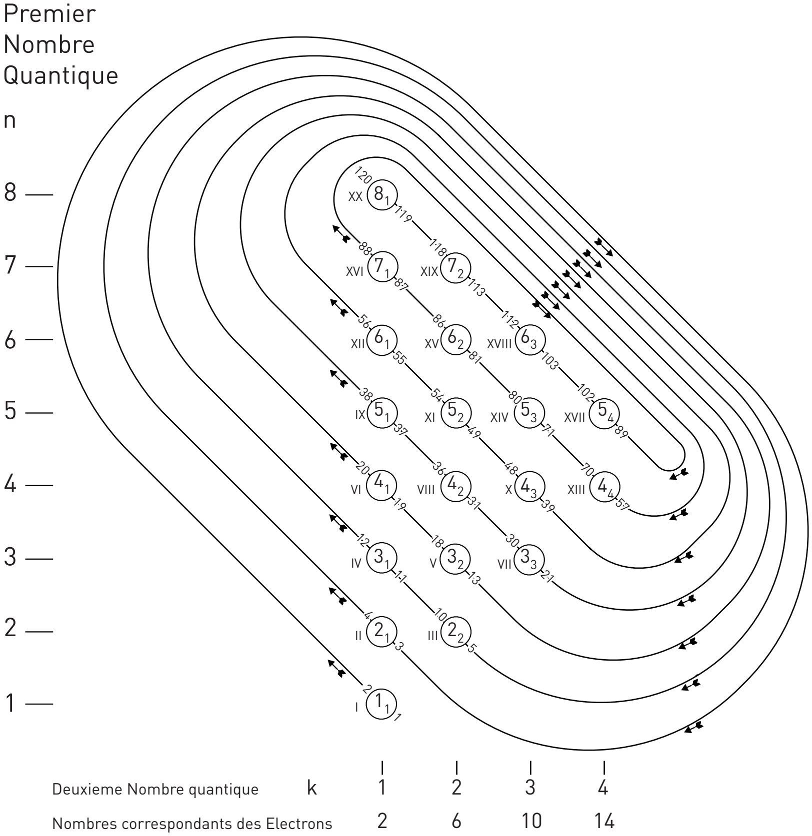


FIGURE 13.6 Graph produced by Janet exhibiting the sequential building-up of electronic shells in terms of the quantum numbers *n* and *k*. [From: C. Janet. *Concordance de l'Arrangement Quantique de Base, des Electrons Planetaires des Atomes, avec la Classification Scalariforme, Hélicoïdale, des Eléments Chimiques.* 1930. Reprinted with permission of the *Société Académique de l'Oise*, Beauvais, France. Courtesy of Loïc Casson.]

his LSPT without taking into account the old quantum theory of Bohr, Edmund C. Stoner, and Arnold Sommerfeld. One year later, however, Janet did notice the perfect agreement between his LSPT and Bohr's planetary atomic model.[20] Anteceding Madelung by six years, Janet recognized the importance of the $n + l$ rule, as can be seen from a graph that he produced in 1930 (Figure 13.6).[21] He also adopted the actinide and transactinide classification, which is usually attributed to Glenn Seaborg (1912–1999).

Despite the power and elegance of Janet's scheme, his classifications failed to attract the attention of the scientific community until this century. In recent times, Eric Scerri, Gary Katz, Henry A. Bent, and Valery Tsimmerman have all advocated the use

---

can be found in L. Casson. "Sur le Fil Rouge de Charles Janet. Approche Épistémologique et Biographique," *Mémoires de la Société d'Archéologie, Sciences et Arts du Département de l'Oise* 37 (2015), pp. 163–197.

[20] C. Janet. *Concordance de l'Arrangement Quantique de Base des Électrons Planétaires des Atomes avec la Classification Scalariforme Hélicoïdale des Éléments Chimiques.* Beauvais: Imprimerie Départementale de l'Oise, 1930.

[21] Ibid. This also explains the alternative names *fdps table* or *n + l ordinal table* for the LSTP.

of the LSPT, arguing that it offers certain advantages in comparison with the more conventional medium-long form of the periodic system[22]:

1. The periods in Janet's table are characterized by a constant value of $N = n + l$, and this is without any exception. In the medium-long form, on the other hand, it seems the periods are characterized by a constant value of n, but this rule is violated several times. Examples include the unexpected appearance of the $3d$-block in the fourth period and the sudden emergence of a $4f$-block in the sixth period. The LSPT is therefore in better agreement with the Madelung $(n + l, n)$ rule, and suggests the possibility of elevating $N = n + l$ to a new *quantum number* for the periodic table.
2. By organizing the elements in periods of constant $n + l$ and groups of constant l, m_l, and m_s, the *period doubling* occurs naturally as an essential feature of the periodic system. The pairing of the periods is responsible for the distinctive stepped profile of the LSPT and it leads to the following set of cardinalities:

   $$2 - 2 - 8 - 8 - 18 - 18 - 32 - 32. \tag{13.18}$$

   This should be contrasted with the more artificial sequence of cardinalities in Eq. (13.1) for the traditional seven-period periodic table.[23]
3. The ordering of the blocks is, in a sense, more "natural" in the LSPT than in the conventional long form periodic table because it follows the natural filling of the orbitals (compare f-d-p-s with s-f-d-p). This ordering leads to a gradual change from strong *horizontal relationships* within the lanthanide and actinide series at the left of the chart, to pronounced *vertical relationships* among the main block elements at the right.[24]

A disadvantage, according to many chemists, is the placement of helium among the alkaline earth metals in group IIA. Despite its agreement with quantum mechanics, numerous chemists have advocated against this placement on the basis of empirical data—a conduct referred to as the *tyranny of the chemist.*[25]

13.4.2 Period doubling

In the previous discussion, the *period doubling* emerged as a characteristic feature of the periodic system. This section offers some *empirical evidence* to support this claim. The doubling of the periods leads to important *secondary periodicity phenomena.* Secondary trends in the periodic table have long been recognized and include, among others, the d-block contraction, lanthanide contraction, tetrad effects, gadolinium break, diagonal relationships, Knight's Move, and inert-pair effects.

---

[22] See, for example, E. R. Scerri. "Chemistry, Spectroscopy, and the Question of Reduction." *Journal of Chemical Education* 68.2 (1991), pp. 122–126; G. Katz. "The Periodic Table: An Eight Period Table for the 21st Century." *Chemical Educator* 6 (2001), pp. 324–332; E. R. Scerri. "Some Aspects of the Metaphysics of Chemistry and the Nature of the Elements." *HYLE* 11.1–2 (2005), pp. 127–145; Bent, *New Ideas in Chemistry from Fresh Energy for the Periodic Law.*

[23] Katz, "The Periodic Table: An Eight Period Table for the 21st Century," p. 325.

[24] L. M. Simmons. "A Modification of the Periodic Table." *Journal of Chemical Education* 24.12 (1947), pp. 588–591.

[25] E. R. Scerri. "The Tyranny of the Chemist." *Chemistry International* 28.3 (2006), p. 11.

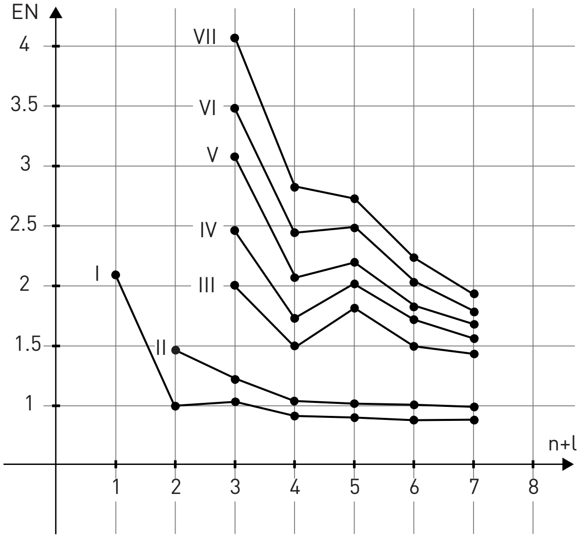


FIGURE 13.7 The electronegativity of the main block elements is plotted for each chemical group as a function of $N = n + l$. The oscillatory behavior is especially pronounced for the group III elements. [Adapted from D. Neubert. "Double Shell Structure of the Periodic System of the Elements." *Zeitschrift für Naturforschung* 25a (1970), p. 212.]

In the case of the period doubling, the manifold of chemical elements splits naturally into two sets, with $n + l$ even or odd. Elements from the same (even or odd) set are, as a result, chemically more similar than elements from different sets. This phenomenon was first brought to the attention of the chemical community by Sanderson in 1952.[26] Numerous examples can be cited:

1. Among the hydrides of the third group, B_2H_6 and Ga_2H_6 are volatile, but $(AlH_3)_x$ is nonvolatile.
2. SiH_4 is more readily oxidized and hydrolyzed than CH_4 and GeH_4.
3. N(III) and As(III) are weaker reducing agents than P(III).

An important implication of the period doubling is that the properties of the congeners within a major group do not vary in a regular and continuous manner, but in a harmonically alternating way. Following Sanderson, this can be easily visualized by plotting some atomic property X of the elements as a function of $N = n + l$ for constant values of l, m_l, and m_s. An oscillatory trend is then observed, which can be seen as a superposition of two continuous functions for $n + l$ odd or even.[27]

Similar plots were produced by Karl-Dietrich Neubert in 1970 for the electronegativity and first ionization energy of the main block elements (Figure 13.7).[28] Neubert referred to this discontinuous behavior as evidence for a *double shell structure* in the

---

[26] R. T. Sanderson. "An Explanation of Chemical Variations within Periodic Major Groups." *Journal of the American Chemical Society* 74.19 (1952), pp. 4792–4794.

[27] See Sanderson, "An Explanation of Chemical Variations within Periodic Major Groups," p. 4793 as well as R. T. Sanderson. Chemical Periodicity. New York: Reinhold, 1960, p. 35.

[28] D. Neubert. "Double Shell Structure of the Periodic System of the Elements." *Zeitschrift für Naturforschung* 25a (1970), pp. 210–217.

periodic table. Not only is there a doubling of the states in the *horizontal* direction resulting from the electronic spin $s = \pm 1/2$, a second doubling seems to occur in the *vertical* direction, which Neubert attributed to a new quantum number, denoted $c = \pm 1/2$.

The doubling of the periods and their empirical ramifications were studied further by Halis Odabaşi (1931–2011).[29] Once again, the ionization potentials for neutral atoms and singly ionized ions were observed to be separated into two groups according to whether $n + l$ is even or odd. For each group, a smooth curve was obtained, and with the help of curve fitting procedures, Odabaşi extrapolated these curves to regions where $Z > 120$.

The above discussion highlights the importance of the period doubling as a characteristic feature of the left-step periodic system. Accordingly, this property will have to be explained if a deeper understanding of the periodic law is ever to be archieved.

13.4.3 *Ab initio* derivation of the $(n+l, n)$ rule

Despite the rising popularity of the $(n+l, n)$ rule, the Madelung rule remains a purely *empirical* (or *lexicographic*) rule. Allen and Knight have called it a "somewhat mysterious algorithm,"[30] whereas Ostrovsky wondered about the "dynamical origin of the sum of principal n and orbital l quantum numbers."[31] In 1969, during the centennial anniversary of Mendeleev's discovery, Per-Olov Löwdin published *Some Comments on the Periodic System of the Elements*, noting how remarkable it was that "in axiomatic quantum theory, the simple $(n+l, n)$ energy rule has not yet been derived from first principles."[32] He concluded his paper by saying,

> It would certainly be worth while to study the $(n+l, n)$ energy rule from first principles, i.e. on the basis of the many-electron Schrödinger equation.[33]

The quest for an *ab initio* derivation of the Madelung rule became known as the *Löwdin challenge*—the "oldest and largest standing problem in quantum chemistry."[34] One must be careful to distinguish this Löwdin challenge from a high-precision calculation on an individual atomic element. The latter is aimed at reproducing the specific atomic properties of that element, whereas the former concerns the understanding of the periodic system as a whole.[35]

The goal of the Löwdin challenge, in other words, is to solve the many-electron Schrödinger equation and obtain the particular linear combination of n and l

---

[29] H. Odabaşi. "Some Evidence about the Dynamical Group SO (4,2). Symmetries of the Periodic Table of Elements." *International Journal of Quantum Chemistry* 7.S7 (1973), pp. 23–33. See also E. U. Condon and H. Odabaşi. "The Order of Electron Shells for Atoms and Ions." In: *Atomic Structure.* Cambridge: Cambridge University Press, 1980, pp. 514–525.

[30] L. C. Allen and E. T. Knight. "The Löwdin Challenge: Origin of the $n + l, n$ (Madelung) Rule for Filling the Orbital Configurations of the Periodic Table." *International Journal of Quantum Chemistry* 90.1 (2002), p. 83.

[31] Ostrovsky, "What and How Physics Contributes to Understanding the Periodic Law," p. 158.

[32] Löwdin, "Some Comments on the Periodic System of the Elements," p. 332

[33] Ibid., p. 334

[34] Allen and Knight, "The Löwdin Challenge", p. 83.

[35] Ostrovsky, "Physical Explanation of the Periodic Table."

quantum numbers as a result.[36] Since Löwdin's plea, many claims to a successful derivation have appeared in the scientific literature, but most have been dismissed—an issue to which we turn in the next section.

13.5 MISAPPLYING THE MADELUNG RULE

13.5.1 The *n+l* blunder

The validity and utility of the $(n+l,n)$ rule has recently been under severe attack by Eugen W. H. Schwarz.[37] He remarked that "it has been deplored from time to time that the Madelung rule has not yet been derived quantum-theoretically" (§13.4.3), but added, "this is of course impossible since this approximate rule of thumb is at variance with too many facts."[38] According to Schwarz "no general convincing theoretical derivation exists because the Madelung rule does not hold in general for the atomic valence shells, and not at all for the core shells" (*vide infra*).[39]

In his opinion, the $(n+l,n)$ rule represents "a nonrealistic and chemically misleading order of atomic orbitals,"[40] which "is of limited value in chemistry."[41] Referring to it as "the $n+l$ blunder,"[42] Schwarz concluded that all "previous wisdom of the periodic system has suffered from the invention of such scientific facts."[43]

Let it be clear at the outset that we do *not* agree with Schwarz's conclusions. Although the $(n+l,n)$ rule might be "at variance with many facts," its explanatory power should not be underestimated, as we intend to explain in what follows. Having said this, we admit that Schwarz has raised a number of important issues that would benefit from closer scrutiny, and which shows the need for a correct *interpretation* of the $(n+l,n)$ rule.

13.5.2 The concept of an element

Before offering two possible interpretations of the $(n+l,n)$ rule in §13.5.3, the aim of this section is to show that the applicability of the Madelung rule greatly depends on the definition given to the concept of a chemical element. A distinction should

---

[36] Ostrovsky, "What and How Physics Contributes to Understanding the Periodic Law," p. 158.

[37] S. G. Wang and W. H. E. Schwarz. "Icon of Chemistry: The Periodic System of Chemical Elements in the New Century." *Angewandte Chemie International Edition* 48 (2009), pp. 3404–3415; W. H. E. Schwarz and S. G. Wang. "Some Solved Problems of the Periodic System of Chemical Elements." *International Journal of Quantum Chemistry* 110.8 (2010), pp. 1455–1465; W. H. E. Schwarz and R. L. Rich. "Theoretical Basis and Correct Explanation of the Periodic System: Review and Update." *Journal of Chemical Education* 87.4 (2010), pp. 435–443; W. H. E. Schwarz. "The Full Story of the Electron Configurations of the Transition Elements." *Journal of Chemical Education* 87.4 (2010), pp. 444–448.

[38] Schwarz and Rich, "Theoretical Basis and Correct Explanation of the Periodic System," p. 441.

[39] Schwarz and Wang, "Some Solved Problems of the Periodic System of Chemical Elements," p. 1461.

[40] Ibid., p. 1457.

[41] Schwarz and Rich, "Theoretical Basis and Correct Explanation of the Periodic System."

[42] Wang and Schwarz, "Icon of Chemistry: The Periodic System of Chemical Elements in the New Century," p. 3412.

[43] Ibid., p. 3404. He conjectured, however, that "since the rule occurs in all textbooks and is absorbed by all students and teachers, it will die out only very slowly." (p. 3412).

be made with respect to the kind of atoms corresponding to a given element. We will distinguish between (1) free and bound atoms, (2) neutral atoms and ionic species, and (3) atoms in their ground and excited state. By way of example, this section will focus on the element nickel (Ni, $Z = 28$).

Free versus bound atoms

Ni^0 as a *free atom* has a ground state electronic configuration [Ar] $3d^8 4s^2$, whereas the dominant configuration for Ni^0 in nickel metal (as a *bounded atom*) or in tetracarbonyl nickel $Ni(CO)_4$ is [Ar] $3d^{10}$.[44] Manifestly, free atoms and bound atoms are very different. The $(n+l, n)$ rule correctly determines the configuration for Ni as a free atom, but does *not* apply to the configuration of Ni as a bound atom.

The origin of these dissimilar configurations can be traced back to a *destabilization* of the $(n+1)\,s$ atomic orbital (AO) in compounds. An $(n+1)\,s$ AO is more extended spatially than an *nd*-orbital. When the atom is surrounded by other atoms in a chemical compound, the $(n+1)\,s$ AO overlaps with the AOs of the other atoms and becomes destabilized as a result of the Pauli exclusion principle. As a consequence, $nd \ll (n+1)s$ energetically, which explains the pure $d^g s^0$ configuration for transition metals in compounds (where g is the group number).[45]

Neutral atoms versus charged ions

The ground state configurations for all free transition metals M^i $(i = 0, +1, +2)$ from the $3d$-to $5d$-block are given in Table 13.3.[46] Among the neutral transition metal atoms M^0, only *one*, palladium (Pd), has a $d^g s^0$ configuration with empty s AO (where g is the group number). Nearly *half* the transition metal cations M^+, on the other hand, have pure d ground states $d^{g-1} s^0$ with empty s shell; and among the higher charged cations M^{i+} $(i > 1)$, practically *all* have ground state configurations $d^{g-i} s^0$ with empty s AOs.

This indicates that, for higher ionic charges (and thus larger Z_{eff}), $nd < (n+1)s$ energetically. This is in agreement with the hydrogenic (n, l) rule, but is in stark contrast with the Madelung $(n+l, n)$ rule. The Madelung rule therefore applies only to neutral atoms whereas, for ions, different filling rules apply.[47]

Ground versus excited states

Finally, it must be emphasized that all traditional periodic systems represent the electronic configurations of the elements in their (stable) *ground state*, and not in a (metastable) *excited state*. Based on this, and the last two sections, we can conclude that the validity of the $(n+l, n)$ rule is limited to *ground state* electronic configurations of *neutral* atoms as *free* gas-phase species in a vacuum. (Although most interesting,

---

[44] For this reason, transition metal chemistry is often termed "pure d–shell chemistry." See also G. Frenking and N. Fröhlich. "The Nature of the Bonding in Transition-Metal Compounds." *Chemical Reviews* 100.2 (2000), pp. 717–774, §III.

[45] Wang and Schwarz, "Some Solved Problems of the Periodic System of Chemical Elements," p. 3406.

[46] See also C. E. Moore. *Atomic Energy Levels as Derived from the Analyses of Optical Spectra.* Washington, DC: US. Government Printing Office, 1949, 1952, 1958.

[47] See Goudsmit and Richards, "The Order of Electron Shells in Ionized Atoms."

Table 13.3 Ground state electronic configurations for all free transition metals M^i from the $3d$-to $5d$-block, with $i = 0, +1, +2$. [Data obtained from the National Institute of Standards and Technology (NIST) Atomic Spectra Database. Available at: http://physics.nist.gov/asd]

M	M^0	M^+	M^{2+}	M	M^0	M^+	M^{2+}	M	M^0	M^+	M^{2+}
Sc	[Ar] $3d4s^2$	[Ar] $3d4s$	[Ar] $3d$	Y	[Kr] $4d5s^2$	[Kr] $5s^2$	[Kr] $4d$	Lu	[Xe] $4f^{14}5d6s^2$	[Xe] $4f^{14}6s^2$	[Xe] $4f^{14}6s$
Ti	[Ar] $3d^24s^2$	[Ar] $3d^24s$	[Ar] $3d^2$	Zr	[Kr] $4d^25s^2$	[Kr] $4d^25s$	[Kr] $4d^2$	Hf	[Xe] $4f^{14}5d^26s^2$	[Xe] $4f^{14}5d6s^2$	[Xe] $4f^{14}5d^2$
V	[Ar] $3d^34s^2$	[Ar] $3d^4$	[Ar] $3d^3$	Nb	[Kr] $4d^45s$	[Kr] $4d^4$	[Kr] $4d^3$	Ta	[Xe] $4f^{14}5d^36s^2$	[Xe] $4f^{14}5d^36s$	[Xe] $4f^{14}5d^3$
Cr	[Ar] $3d^54s$	[Ar] $3d^5$	[Ar] $3d^4$	Mo	[Kr] $4d^55s$	[Kr] $4d^5$	[Kr] $4d^4$	W	[Xe] $4f^{14}5d^46s^2$	[Xe] $4f^{14}5d^46s$	[Xe] $4f^{14}5d^4$
Mn	[Ar] $3d^54s^2$	[Ar] $3d^54s$	[Ar] $3d^5$	Tc	[Kr] $4d^65s$	[Kr] $4d^55s$	[Kr] $4d^5$	Re	[Xe] $4f^{14}5d^56s^2$	[Xe] $4f^{14}5d^56s$	[Xe] $4f^{14}5d^5$
Fe	[Ar] $3d^64s^2$	[Ar] $3d^64s$	[Ar] $3d^6$	Ru	[Kr] $4d^75s$	[Kr] $4d^7$	[Kr] $4d^6$	Os	[Xe] $4f^{14}5d^66s^2$	[Xe] $4f^{14}5d^66s$	[Xe] $4f^{14}5d^56s$
Co	[Ar] $3d^74s^2$	[Ar] $3d^8$	[Ar] $3d^7$	Rh	[Kr] $4d^85s$	[Kr] $4d^8$	[Kr] $4d^7$	Ir	[Xe] $4f^{14}5d^76s^2$	[Xe] $4f^{14}5d^76s$	[Xe] $4f^{14}5d^7$
Ni	[Ar] $3d^84s^2$	[Ar] $3d^9$	[Ar] $3d^8$	Pd	[Kr] $4d^{10}$	[Kr] $4d^9$	[Kr] $4d^8$	Pt	[Xe] $4f^{14}5d^96s$	[Xe] $4f^{14}5d^9$	[Xe] $4f^{14}5d^8$
Cu	[Ar] $3d^{10}4s$	[Ar] $3d^{10}$	[Ar] $3d^9$	Ag	[Kr] $4d^{10}5s$	[Kr] $4d^{10}$	[Kr] $4d^9$	Au	[Xe] $4f^{14}5d^{10}6s$	[Xe] $4f^{14}5d^{10}$	[Xe] $4f^{14}5d^9$
Zn	[Ar] $3d^{10}4s^2$	[Ar] $3d^{10}4s$	[Ar] $3d^{10}$	Cd	[Kr] $4d^{10}5s^2$	[Kr] $4d^{10}5s$	[Kr] $4d^{10}$	Hg	[Xe] $4f^{14}5d^{10}6s^2$	[Xe] $4f^{14}5d^{10}6s$	[Xe] $4f^{14}5d^{10}$

we will not pursue the naturally arising question of whether free atoms in a vacuum represent the elements as basic substances.) Conversely, in some cases, the ground state occupation may differ from the Hartree-Fock configuration average as a result of electron correlations beyond the self-consistent field approximation. This effect is most pronounced near half-filled or closed *d*-shells. As an example, for Cr, the lowest average energy configuration is $3d^4 4s^2$. Nevertheless, the ground state is the ^7S multiplet based on $3d^5 4s^1$ as a result of its favorable spin pairing energy.

13.5.3 Two interpretations of the $(n+l, n)$ rule

The $(n+l, n)$ rule as the Aufbau principle

Among the most common interpretations of the $(n+l, n)$ rule is the assumption that the Madelung rule affords us with a correct ordering of the different *nl* states in terms of increasing energy

$$\varepsilon(1s) \ll \varepsilon(2s) < \varepsilon(2p) \ll \varepsilon(3s) < \varepsilon(3p) \ll \varepsilon(4s) < \varepsilon(3d) < \varepsilon(4p) \ll \ldots. \quad (13.19)$$

This order is assumed to be *fixed* and *universal*.[48] That is, the energetic order $\varepsilon(4s) < \varepsilon(3d)$, for example, is asserted to hold at all times. We will challenge this interpretation (1) by a brief consideration of the famous $3d/4s$ issue and (2) by putting the Madelung sequence in the broader context of AO sequences.

The 3d/4s issue. Let us verify the validity of Eq. (13.19) by building a scandium (Sc) atom from scratch. We start with a bare $^{45}_{21}\mathrm{Sc}$ nucleus, consisting of twenty-one protons (p^+) and twenty-four neutrons (n^0), and proceed by adding a total of twenty-one electrons (e^-) to form a neutral Sc atom. This addition happens in a stepwise manner (one electron at a time); each electron is accommodated in an *nl*-orbital according to the Aufbau principle (Eq. (13.19)), while keeping Pauli's exclusion principle in mind.

After the addition of eighteen electrons, a Sc^{3+} ion is formed with an argon configuration $[\mathrm{Ar}] = [\mathrm{Ne}]\ 3s^2 3p^6$. The next electron is placed in the $4s$-orbital, in accordance, with Eq. (13.19), to yield a Sc^{2+} cation with configuration $[\mathrm{Ar}]\ 4s^1$. The $4s$-shell is then completed by the addition of the twentieth electron, resulting in a Sc^+ ion with $[\mathrm{Ar}]\ 4s^2$ configuration. The twenty-first and last electron ends up in a $3d$-orbital, yielding a neutral Sc atom with configuration $[\mathrm{Ar}]\ 3d^1 4s^2$. This Aufbau process can be summarized as follows:

$$\mathrm{Sc}^{3+} \to \mathrm{Sc}^{2+}\,(4s^1) \to \mathrm{Sc}^{+}\,(4s^2) \to \mathrm{Sc}\,(3d^1 4s^2). \quad (13.20)$$

This, however, is contradicted by the experimental data (Table 13.3), which shows that

$$\mathrm{Sc}^{3+} \to \mathrm{Sc}^{2+}\,(3d^1) \to \mathrm{Sc}^{+}\,(3d^1 4s^1) \to \mathrm{Sc}\,(3d^1 4s^2). \quad (13.21)$$

Here, the first valence electron goes into the energetically lower $3d$-shell, contrary to the (13.19) ordering. A second difference concerns the $3d^1 4s^1$ configuration of Sc^+, which is explained by the fact that the $3d$ AO is spatially more compact than the $4s$ AO. The electron–electron repulsion is therefore much higher in the $3d$-orbital than in the larger $4s$-shell, and this favors the $3d^1 4s^1$ configuration energetically, with a

---

[48] The universality of Eq. (13.19) has already been disputed in §13.5.2.

second valence electron being placed in the higher 4*s* AO, where it experiences smaller repulsion.[49]

The common textbook dictum "4*s* is occupied first, but 4*s* is also ionized first," which, according to Schwarz, "seems to violate the principle of *microscopic reversibility*,"[50] should thus be corrected by saying that in a transition metal atom, 3*d* is occupied *before* 4*s*, wherefore 4*s* is logically easier to ionize.[51] It bears repeating that this is at variance with the energy sequence in Eq. (13.19), and it challenges the previously mentioned interpretaton of the $(n+l, n)$ rule as an Aufbau principle.

Different AO sequences. According to Schwarz, this misinterpretation is a result of the fact that "most chemistry textbooks teach one single AO sequence, corresponding to the Madelung rule," whereas at least five sequences are known to exist.[52] The Fock (n) sequence (without screening by other electrons)

$$1s \ll 2s = 2p \ll 3s = 3p = 3d \ll 4s = 4p = 4d = 4f \ll 5s = 5p = 5d = \ldots, \quad (13.22)$$

and the Madelung $(n+l, n)$ sequence (with a strongly screened nucleus)

$$1s \ll 2s < 2p \ll 3s < 3p \ll 4s < 3d < 4p \ll 5s < 4d < 5p \ll 6s < 4f < \ldots, \quad (13.23)$$

represent two *limiting cases*, with the (actual) realistic sequences lying in between these two extremes as intermediates.[53] The AO sequence for the *inner atomic core shells*, for instance, follows the hydrogenic (n, l) rule:

$$1s \ll 2s < 2p \ll 3s < 3p < 3d \ll 4s < 4p < 4d < 4f \ll 5s < 5p < 5d < \ldots, \quad (13.24)$$

---

[49] Wang and Schwarz, "Icon of Chemistry: The Periodic System of Chemical Elements in the New Century," p. 3407.

[50] Ibid., p. 3406.

[51] See also S. G. Wang, Y. X. Qiu, H. Fang, and W. H. E. Schwarz. "The Challenge of the So-Called Electron Configurations of the Transition Metals." Chemistry: A European Journal 12.15 (2006), pp. 4101–4114; and J. B. Mann, T. L. Meek, E. T. Knight, J. F. Capitani, and L. C. Allen, "Configuration Energies of the *d*-Block Elements." *Journal of the American Chemical Society* 122.21 (2000), pp. 5132–5137. The seemingly preferential occupation of the 4s-level before the 3d-level, according to Eq. (13.19), has been the source for numerous heated discussions in the scientific literature. See F. L. Pilar. "4s Is Always Above 3d! or, How to Tell the Orbitals from the Wave Functions." *Journal of Chemical Education* 55.1 (1978), pp. 2–6; T. S. Carlton. "4s Sometimes Is below 3d." *Journal of Chemical Education* 56.11 (1979), p. 767; F. L. Pilar. "4s Sometimes Is below 3d (the Author Replies)." *Journal of Chemical Education* 56.11 (1979), p. 767; E. R. Scerri. "Transition Metal Configurations and Limitations of the Orbital Approximation." *Journal of Chemical Education* 66.6 (1989), pp. 481–483; L. G. Vanquickenborne, K. Pierloot, and D. Devoghel. "Electronic Configuration and Orbital Energies: The 3d-4s Problem." *Inorganic Chemistry* 28.10 (1989), pp. 1805–1813; L. G. Vanquickenborne, K. Pierloot, and D. Devoghel. "Transition Metals and the Aufbau Principle." *Journal of Chemical Education* 71.6 (1994), pp. 469–471; M. P. Melrose and E. R. Scerri. "Why the 4s Orbital Is Occupied before the 3d." *Journal of Chemical Education* 73.6 (1996), pp. 498–503; J. L. Bills. "Why the 4s Orbital Is Occupied before the 3d." *Journal of Chemical Education* 74.6 (1997), p. 616; M. P. Melrose and E. R. Scerri. "Why the 4s Orbital Is Occupied before the 3d (the Authors Reply)." *Journal of Chemical Education* 74.6 (1997), p. 616; J. L. Bills. "Experimental 4s and 3d Energies in Atomic Ground States." *Journal of Chemical Education* 75.5 (1998), pp. 589–593.

[52] Wang and Schwarz, "Icon of Chemistry: The Periodic System of Chemical Elements in the New Century," p. 3407. See also Schwarz and Wang, "Some Solved Problems of the Periodic System of Chemical Elements."

[53] Wang and Schwarz, "Icon of Chemistry: The Periodic System of Chemical Elements in the New Century," p. 3409.

where, as a general rule, $ns < np \ll nd \lll nf$ as a result of a difference in angular momentum $L = \sqrt{l(l+1)}\hbar$ and shielding/penetrating power (see §13.3.2). This has been verified experimentally via X-ray spectroscopy and quantum mechanical calculations.

For the *outer valence shells*, these shifts become even more pronounced, as determined by ultraviolet/visible spectroscopy, resulting in the following AO sequence:

$$1s \ll 2s < 2p \ll 3s < 3p \ll 3d < 4s < 4p \ll 4d < 5s < 5p \ll 4f < 5d < \ldots. \quad (13.25)$$

Notice how, compared with (Eq. (13.24)), the $3d$-level approaches $4s$, and how the $4f$-orbitals are shifted above the $5s$ and $5p$ AOs.

These screening effects are largest for the alkali and alkaline earth metals of group 1 and 2 because of the large ratio of screening core electrons to valence electrons. This yields a slightly modified AO sequence for the valence electrons of the electropositive elements:

$$1s \ll 2s < 2p \ll 3s < 3p \ll 4s < 3d < 4p \ll 5s < 4d < 5p \ll 6s < 4f \ldots, \quad (13.26)$$

where the $(n+1)\,s$ AOs move down below the nd AOs.[54] The Madelung sequence in Eq. (13.23) is most similar to this last sequence. This induced Schwarz to downgrade the universality of the Madelung rule as applying "only to the first two groups of the periodic system."[55]

Conclusion. The sheer variety of AO sequences and the induced *configurational reorganizations* when adding electrons to a bare nucleus refute the simplistic interpretation of the $(n+l,n)$ rule as a fixed energy-ordering principle. The change in the order of nl-orbitals when moving from the inner core to the outer valence regions of the atom reduces the Madelung sequence to an overly simplistic summary of the spectroscopic results and calls for an alternative interpretation (*vide infra*).

The (*n+l,n*) rule as a tool for predicting electron configurations

Although the $(n+l,n)$ rule fails to accurately describe the entire *building-up* process of an atom (as exemplified by the Sc case presented earlier), it does seem to correctly predict the *overall* (final) electronic ground state configuration of a neutral atom. Stated somewhat differently, the $(n+l,n)$ rule establishes in which nl spin orbital the *differentiating electron* should go.[56]

Yet, many exceptions are known to exist that undermine this second interpretation of the Madelung rule. Terry L. Meek and Leland C. Allen have listed a total of nineteen elements (ten d-block and nine f-block elements) exhibiting *anomalous configurations* that differ from those predicted by the $(n+l,n)$ rule (Table 13.4).[57] It is interesting to note that the configurations of the ten d-block elements present a deviation from the first $(n+l)$ part of the Madelung rule, whereas the configurations of the nine f-block elements violate the second (n) part. Seventeen of these "non-Madelung"

---

[54] Ibid., p. 3408.

[55] Schwarz and Rich, "Theoretical Basis and Correct Explanation of the Periodic System: Review and Update", p. 436.

[56] The *differentiating electron* is defined as that electron by which two successive elements differ.

[57] T. L. Meek and L. C. Allen. "Configuration Irregularities: Deviations from the Madelung Rule and Inversion of Orbital Energy Levels." *Chemical Physics Letters* 362.5–6 (2002), pp. 362–364.

Table 13.4 Ground-state electronic configurations for nineteen elements exhibiting anomalous configurations that do not obey the $(n+l, n)$ rule. [Data obtained from the National Institute of Standards and Technology (NIST) Atomic Spectra Database. Available at: http://physics.nist.gov/asd]

Element	*Predicted Madelung ground state configuration*	*Experimentally obtained ground state configuration*
Cr	[Ar] $3d^4 4s^2$	[Ar] $3d^5 4s^1$
Cu	[Ar] $3d^9 4s^2$	[Ar] $3d^{10} 4s^1$
Nb	[Kr] $4d^3 5s^2$	[Kr] $4d^4 5s^1$
Mo	[Kr] $4d^4 5s^2$	[Kr] $4d^5 5s^1$
Ru	[Kr] $4d^6 5s^2$	[Kr] $4d^7 5s^1$
Rh	[Kr] $4d^7 5s^2$	[Kr] $4d^8 5s^1$
Pd	[Kr] $4d^8 5s^2$	[Kr] $4d^{10}$
Ag	[Kr] $4d^9 5s^2$	[Kr] $4d^{10} 5s^1$
La	[Xe] $4f^1 6s^2$	[Xe] $5d^1 6s^2$
Ce	[Xe] $4f^2 6s^2$	[Xe] $4f^1 5d^1 6s^2$
Gd	[Xe] $4f^8 6s^2$	[Xe] $4f^7 5d^1 6s^2$
Pt	[Xe] $4f^{14} 5d^8 6s^2$	[Xe] $4f^{14} 5d^9 6s^1$
Au	[Xe] $4f^{14} 5d^9 6s^2$	[Xe] $4f^{14} 5d^{10} 6s^1$
Ac	[Rn] $5f^1 7s^2$	[Rn] $6d^1 7s^2$
Th	[Rn] $5f^2 7s^2$	[Rn] $6d^2 7s^2$
Pa	[Rn] $5f^3 7s^2$	[Rn] $5f^2 6d^1 7s^2$
U	[Rn] $5f^4 7s^2$	[Rn] $5f^3 6d^1 7s^2$
Np	[Rn] $5f^5 7s^2$	[Rn] $5f^4 6d^1 7s^2$
Cm	[Rn] $5f^8 7s^2$	[Rn] $5f^7 6d^1 7s^2$

configurations differ from the predicted configurations by a single electron; whereas palladium (Pd) and thorium (Th) differ by two electrons (see also Figure 13.8).

Being subject to no less than ninenteen exceptions, many an author has called into doubt the lawfulness of the $(n+l, n)$ rule. J. F. Ogilvie, for instance, downgraded the $(n+l, n)$ rule to an "ad hoc rule … of quite limited utility."[58] Benjamin Carroll and Alexander Lehrman, on the other hand, considered these exceptions "relatively unimportant to the chemist."[59] After all, the nineteen "anomalous atoms" have *excited states* near the ground state that *do* satisfy the $(n+l, n)$ rule.

In this regard, Yury N. Demkov and Valentin N. Ostrovsky offered an interesting comparison of the hydrogenic (n, l) and Madelung $(n+l, n)$ rules (Figures 13.8 and 13.9).[60] An exact implementation of one of these rules would result in a perfectly triangular diagram. The $(n+l, n)$ filling rule (Figure 13.8) shows some minor

---

[58] J. F. Ogilvie. "The Nature of the Chemical Bond–1990: There Are No Such Things as Orbitals!" *Journal of Chemical Education* 67.4 (1990), p. 282.

[59] B. Carroll. "The Electron Configuration of the Ground State of the Elements." *Journal of Chemical Education* 25.12 (1948), p. 662.

[60] Yu. N. Demkov and V. N. Ostrovsky. "$n+l$ Filling Rule in the Periodic System and Focusing Potentials." *Soviet Physics JETP* 35.1 (1972), pp. 66–69. Original Russian publication: Yu. N. Demkov and V. N. Ostrovsky. "$n+l$ Filling Rule in the Periodic System and Focusing Potentials."

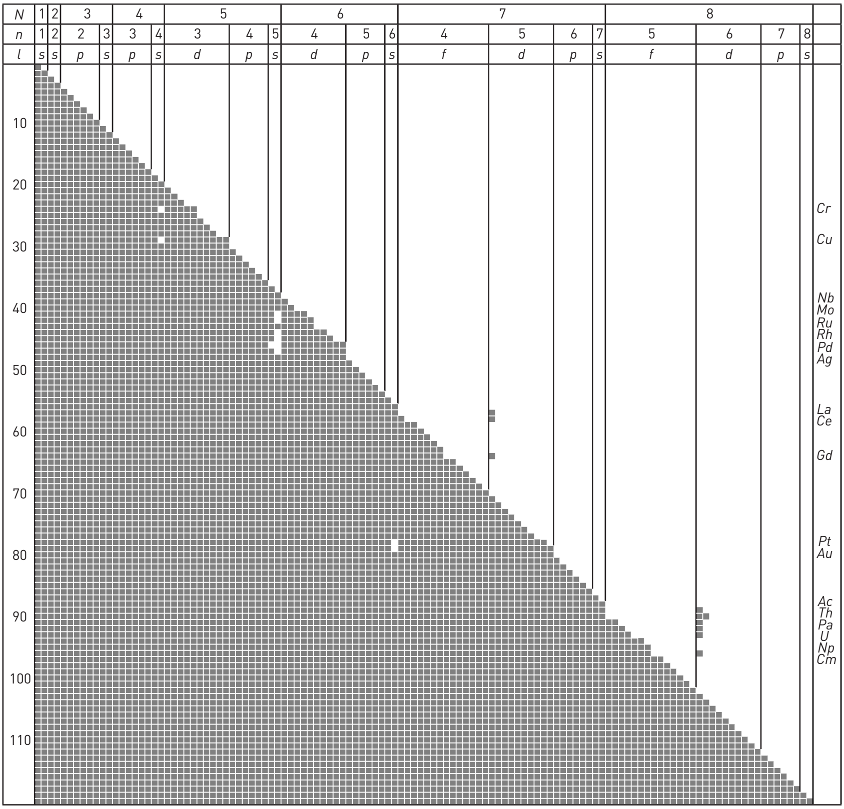


FIGURE 13.8 Madelung filling scheme illustrating the empirical ground state configurations of the chemical elements (with $Z = 1 \rightarrow 120$). The vertical axis represents the atomic number Z; the one-electron nl orbitals are ordered along the horizontal axis according to the $(n+l, n)$ rule. Blackened areas correspond to filled states. An ideal realization of the $(n+l, n)$ rule results in a perfectly triangular diagram. Minor deviations from the triangular shape are observed, corresponging to the nineteen anomalous configurations of §13.5.3. These are also indicated in the rightmost column. [Adapted from Yu. N. Demkov and V. N. Ostrovsky. "n+l Filling Rule in the Periodic System and Focusing Potentials." In: Soviet Physics JETP 35.1 (1972), p. 66.]

deviations from the triangular shape that correspond to the nineteen anomalous configurations of Table 13.4. Yet, the overall triangular shape is clearly visible and, as Ostrovsky noted, "hardly any objective observer would deny that the rule describes well the major trend and thus organizes our knowledge about electron configurations in atoms."[61]

This becomes particularly evident when we compare the $(n+l, n)$ filling scheme with the alternative hydrogenic (n, l) scheme (Figure 13.9). Although both rules work just as fine for low atomic numbers ($Z < 19$), a complete breakdown of the (n, l) rule is observed for higher ordinal numbers ($Z \geq 19$). Unmistakably, then, the $(n+l, n)$ rule quite successfully describes the *electronic structure* of the periodic system (notwithstanding the few imperfections), whereas the (n, l) rule fails over the entire line.

Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki (ZhÉTF) 62 (1972), pp. 125 –132. See also Ostrovsky, "The Periodic Table and Quantum Physics," pp. 342–345.

[61] Ibid., p. 342.

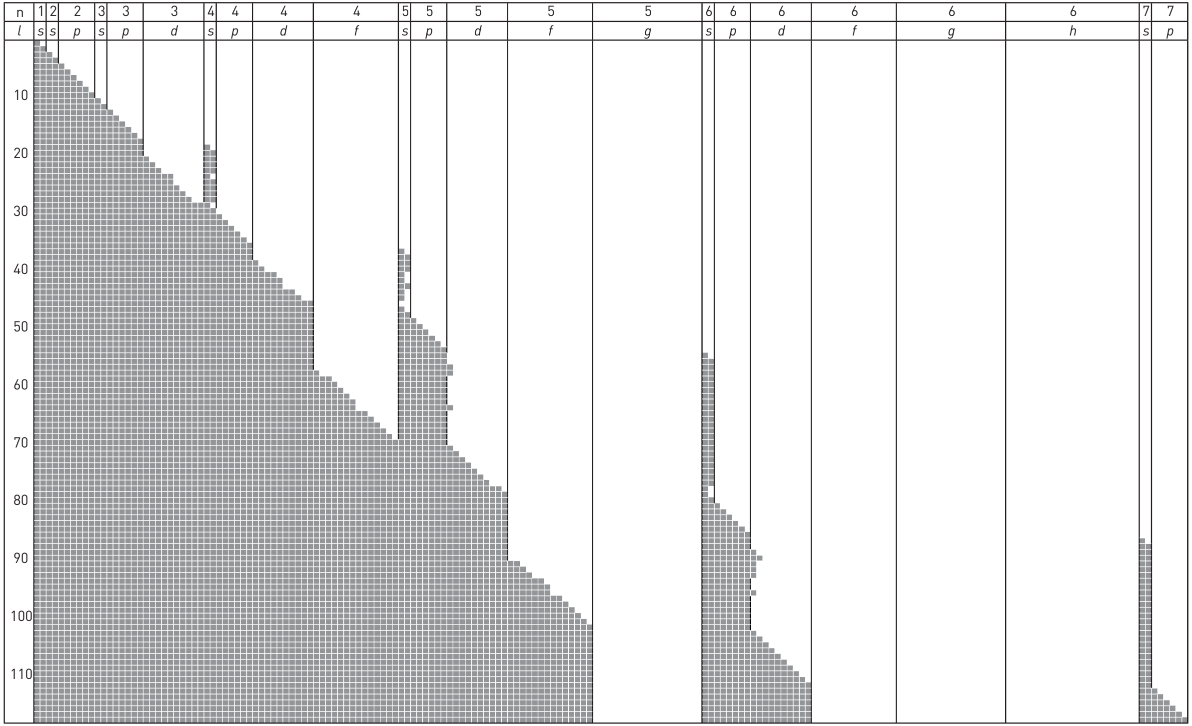


FIGURE 13.9 Hydrogenic (n, l) filling scheme. In this case, the grouping of the *nl* states is carried out in accordance with the hydrogenic (n, l) rule. Although the (n, l) rule works perfectly fine for low atomic numbers ($Z < 19$), a complete breakdown occurs for higher ordinal numbers ($Z \geq 19$). [Adapted from Yu. N. Demkov and V. N. Ostrovsky. "n + l Filling Rule in the Periodic System and Focusing Potentials." In: Soviet Physics JETP 35.1 (1972), p. 67.]

13.6 GROUP THEORY AND THE PERIODIC SYSTEM

In this section, a *group theoretical articulation* of the *global group structure* of the periodic system will be suggested. This raises many questions: In which sense can we speak of the *symmetries* of the periodic system? How are these symmetries described by a *group*? And how does one get from these symmetry groups to an actual *classification* of the elements and a rationalization of the Madelung $(n+l,n)$ rule and period doubling? Answering these questions will be the aim of the remainder of this chapter.

Two major challenges lie ahead if we want to rationalize the degeneracy patterns in the left-step periodic system: (1) provide a symmetry-based interpretation of the *period doubling* and (2) reinterpret the *Madelung* $(n+l,n)$ *rule* in a group theoretical way.

The group theoretical approach to the periodic table has been criticized by numerous authors. According to Pekka Pyykkö, there is "no deep group-theoretical principle" behind the apparent symmetry of the periodic system.[62] Schwarz argued that this approach reduces to the "design of a symmetry group that just fits to the Madelung [atomic orbital] order."[63] Finally, Scerri and colleagues regretted that this approach yields explanations that are "based on symmetry groups, which, by themselves, have to be chosen *empirically*." This "defeats the original purpose" to obtain a "*fundamental* explanation of the periodic table, that is not based on empirical evidence, but on first principles [emphasis in original]."[64]

These criticisms will be countered in the next section, §13.6.1, by making a clear distinction between the *elementary particle approach* (EPA), which is phenomenological in nature, and the *atomic physics approach* (APA), which is more "fundamental" in the above sense. A description of the methodology and philosophy of the group theoretical approach will then be given in §13.6.2, and we will end the chapter in §13.7 with an overview of the literature on this topic.

13.6.1 APA versus EPA

An important distinction can be made between two types of group theoretical approaches.[65] So far, the only quantum systems under study have been the hydrogen atom and the harmonic oscillator, with Hamiltonians that are known exactly. Historically, when group theory started to be applied in atomic and molecular physics,

---

[62] P. Pyykkö. "Relativistic Effects on Periodic Trends." *The Effects of Relativity in Atoms, Molecules and the Solid-State.* Eds. S. Wilson, I. P. Grant, and B. L. Gyorffy. New York: Plenum, 1991, p. 3. See also P. Pyykkö and Y. Zhao. "The Elements of Flatland: Hartree-Fock Atomic Ground States in Two Dimensions for $Z = 1$–24." *International Journal of Quantum Chemistry* 40.4 (1991), pp. 527–544; P. Pyykkö. "A Note on Nodal Structures, Partial Screening, and Periodic Trends among Alkali Metals and Alkaline Earths." *International Journal of Quantum Chemistry* 85.1 (2001), pp. 18–21.

[63] Schwarz and Wang, "Some Solved Problems of the Periodic System of Chemical Elements," p. 1461.

[64] E. Scerri, V. Kreinovich, P. Wojciechowski, and R. R. Yager. "Ordinal Explanation of the Periodic System of Chemical Elements." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6.4 (1998), p. 6.

[65] V. N. Ostrovsky. "Group Theory Applied to the Periodic Table of the Elements." In: *The Mathematics of the Periodic Table.* Eds. D. H. Rouvray and R. B. King. New York: Nova Science Publishers, 2006, pp. 268–269. See also V. N. Ostrovsky. "Group Theory and Periodic System of Elements", p. 196, and Ostrovsky, "What and How Physics Contributes to Understanding the Periodic Law," p. 165.

this was typically the case. Following Ostrovsky's terminology, let us refer to this as the *atomic physics approach*, or APA.

When it comes to the periodic system, however, it is much harder to construct a Hamiltonian, let alone to study its symmetries. This situation could be compared with the status of elementary particle physics during the late 1950s. Because the internal dynamics of the hadrons were shrouded in mist, an exact knowledge of the Hamiltonian system was unavailable.[66] This rendered it naturally impossible to study the symmetry group that would leave the Hamiltonian invariant. Instead, a *phenomenological* approach was adopted; symmetry groups were no longer derived from first principles, but were simply *postulated* on the basis of the available *empirical* data (such as mass properties, conservations laws, and so on). This approach is referred to as the *elementary particle approach*, or EPA, and as we saw in Chapter 7, it ultimately led to a successful classification of the hadrons in the *eightfold way*.[67]

A similar phenomenological approach can be adopted in the study of the periodic table. The structure of the left-step periodic system will serve as *empirical input*; that is, the set of period lengths in Eq. (13.17) will be assumed to represent the dimensions of the unirreps of a particular symmetry group, which will have to be uncovered in the final chapter. A more in-depth study of the internal dynamics of many-electron systems would of course yield additional insights into the physical origin of the observed symmetries in the periodic table.

13.6.2 Group theoretical classification of the elements

In the previous chapters (9–12), the symmetries of the hydrogen atom were studied in great detail. Starting from the spherical SO(3) group, a traditional *symmetry ascent process* was followed that culminated in a derivation of the noninvariance group SO(4,2).[68]

The reason so much space was devoted to the hydrogenic symmetries is because the SO(4,2) group provides a window to the whole of chemistry. It not only establishes from first principles the units of length and energy for atomic phenomena, but it also provides an excellent starting point for a group-theoretical study of the periodic table, based on integer quantum numbers. Of course, whereas the degeneracies of the SO(4) hydrogen shells run as a sequence of squares:

$$1, 4, 9, 16, \ldots, \tag{13.27}$$

the actual magic numbers of the periodic table run as a sequence of double squares. But, this difference can be easily explained *a posteriori* by the introduction of the

---

[66] The theory of strong interactions, *quantum chromodynamics*, was developed only after the discovery of quarks.

[67] M. Gell-Mann and Y. Ne'eman. *The Eightfold Way*. Boulder, CO: Westview Press, 2000.

[68] A second approach to the SO(4,2) group has been proposed by Kibler and Négadi, who followed a *symmetry descent process* starting from the real symplectic group Sp(8). See M. R. Kibler and T. Negadi. "On the Connection between the Hydrogen Atom and the Harmonic Oscillator: The Continuum Case." *Journal of Physics A: Mathematical and General* 16.18 (1983), pp. 4265–4268. M. R. Kibler. "Connection between the Hydrogen Atom and the Harmonic Oscillator: The Zero-Energy Case." *Physical Review A* 29.5 (1984), pp. 2891–2894. See also M. R. Kibler. "On the Use of the Group SO(4,2) in Atomic and Molecular Physics." *Molecular Physics* 102.11–12 (2004), pp. 1221–1229.

additional spin quantum numbers, which are responsible for the doubling of the occupation numbers (see Chapter 8). More troublesome are the actual lengths of the periods, which show a further doubling. Clearly, as far as the periodic structure is concerned, the entrance into the promised land will not be immediate.

The principles of particle classification will be applied to the system of chemical elements. The following sections offer a short description of (1) the different *states* α, β, γ, and so on of (2) the *system* $\mathscr{S}$ and (3) the *symmetry group* $\mathcal{G}$, with (4) its chain of *subgroups*. A refinement of the current approach will then be proposed to deal with *quantum systems* (rather than with *classical systems*). Finally, a justification will be offered for the phenomenological approach that will be taken.

The concept of an atom

In the elementary particle approach that will be adopted in the following pages, the *concept of an atom* differs significantly from the commonplace conception of an atom as defined in mainstream chemistry or quantum mechanics. Its most distinguishing properties are enumerated here:

1. The atoms of the chemical elements are treated as *structureless particles*; the traditional conception of an atom as consisting of a central nucleus surrounded by a cloud of electrons is thus *excluded* from our consideration. Instead, the atoms are assumed to be *noncomposite*. As Byakov and colleagues emphasized, the group theoretical approach "is *not* a [quantum] theory of electronic shells" [emphasis in original].[69] Any talk of orbitals, quantum numbers, or ground state electronic configurations, is considered *meaningless* and *devoid of explanatory power*, and the internal dynamics of atoms can thus be safely ignored.

   You might wonder whether such an approach is justified. After all, there is no disputing that atoms *do* have a complex structure, so why ignore it? An answer to this question will be formulated in a later section.
2. Instead, the atoms of all the different chemical elements will be considered to be various *states* α, β, γ, and so on, of a single *system* $\mathscr{S}$, which in turn will be treated as a kind of *superparticle*.[70] The system $\mathscr{S}$ has been variously named. Rumer and Fet called $\mathscr{S}$ a *Coulombian system*,[71] or a *superparticle*, whereas Barut preferred the term *atomic matter*.[72] Let us follow Barut's terminology and adopt the term *atomic matter* or *primary matter*. Following a proposition made by

---

[69] V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet *Group-Theoretical Classification of Chemical Elements: I. Physical Foundations*. Moscow: Preprint of the Institute of Theoretical and Experimental Physics, ITEP-26, 1976. p. 3.

[70] It should be emphasized that this notion of a *superparticle* has no relation to the supersymmetric partners of bosons and fermions, the existence of which has been postulated in particle physics as the result of a supersymmetry.

[71] Yu B. Rumer and A. I. Fet. "Gruppa Spin(4) i Tablitza Mendeleeva." *Teoreticheskaya i Matematicheskaya Fizika* 9.2 (1971), pp. 203–210. Translated from the Russian: Yu B. Rumer and A. I. Fet. "The Group Spin (4) and the Mendeleev System." *Theoretical and Mathematical Physics* 9.2 (1971), pp. 1081–1085.

[72] A. O. Barut. "Group Structure of the Periodic System." *The Structure of Matter: Rutherford Centennial Symposium*. Ed. B. G. Wybourne. Christchurch, New Zealand: University of Canterbury Press, 1972, pp. 126–136.

Table 13.5 Comparison of the spectrum generating symmetries in the hydrogen atom and in atomic matter.

Quantum system ($\mathscr{S}$)	↔	*Different states* ($\alpha, \beta, \ldots$)
Hydrogen atom	↔	Ground and excited states
Atomic matter	↔	Chemical elements as states

Wulfman in 1978, this pseudoparticle, with a spectrum that is Barut's atomic supermultiplet, can also be denoted by the name *baruton*.[73]

It is important to contrast this approach with the group theoretical treatment of the hydrogen atom in Chapters 5 through 12, where only *one* atom was studied, and where the different *ground* and *excited states* of that single atom represented the different states of the quantum system (see Table 13.5).

The different states (or elements) will thus be represented by ket vectors $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, and so on, that form a *basis* for an infinite-dimensional Hilbert space $\mathcal{H}$ of the system $\mathscr{S}$.

3. All allotropes (e.g., graphite vs. diamond), isotopes (e.g., ${}^{12}_{6}\mathrm{C}$ vs. ${}^{14}_{6}\mathrm{C}$), atoms, ions (e.g., C vs. C^{4+}), and ground and excited states of an element will be treated as one and the same state of the system $\mathscr{S}$. The atom of a chemically specific element is thus nothing more than the carrier of all possible chemical properties. [74] Its only characteristic properties are the atomic number Z and a set of four "quantum numbers" that label the atom.

Atomic matter

The system $\mathscr{S}$ is treated as a *superparticle* with states that are the different elements. From this point of view, $\mathscr{S}$ resembles the *proto hyle* of the ancient Greek philosophers. Otherwise known as Aristotle's *materia prima*, the proto hyle was considered to be the *prime matter* from which emanated all material manifestations. Being itself without substantial form, the primordial matter could be impressed with different kinds of *form* to yield all the different "stuff" of our universe.[75] In a sense, the proto hyle was the *bearer* of all chemical and physical properties, while being itself devoid of any properties. It had, in other words, the *potentiality* of being imbued with any form, but remained denuded from any property as long as no particular form was *actualized* and a specific compound materialized.[76] In short:

$$\text{Matter } (\textit{hyle}) + \text{form } (\textit{morphe}) \rightarrow (\textit{hylomorphic}) \text{ compound.}$$

---

[73] C. E. Wulfman. "Dynamical Groups in Atomic and Molecular Physics." In: *Recent Advances in Group Theory and Their Application to Spectroscopy*. Ed. J. C. Donini. New York: Plenum Press, 1978, pp. 329–403.

[74] V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet, "Group-Theoretical Classification of Chemical Elements. I. Physical Foundations," pp. 3, 19.

[75] These manifestations were called *hylomorphic compounds*—compounds of matter (*hyle*) and form (*morphe*).

[76] "Prime matter is matter in the unqualified sense." Quoted from J. F. Wippel. "The Distinction between Matter and Form." *The Metaphysical Thought of Thomas Aquinas: From Finite Being to Uncreated Being*. Washington D.C.: Catholic University of America Press, 2000, p. 298. The primordial substance derives its name from the Latin *substantia*, which literally means "that which stands under" (i.e., *underlies*).

We can conceive of the system $\mathscr{S}$ in a similar vein as representing the *primary* (*atomic*) *matter* that contains the potentiality of materializing into every possible element when impressed with the appropriate form. Later on, the system $\mathscr{S}$ will also be called the *baruton* in honor of the Turkish-American theoretical scientist Asim Orhan Barut (1926–1994), who initiated the quest for a group theoretical understanding of the periodic system (§ 13.7.1).

The symmetry group $\mathcal{G}$

In analogy with the example of the triangle in §1.2.2, the system $\mathscr{S}$ is subject to a *symmetry group* $\mathcal{G}$. This group will later be identified with the direct product group SO(4,2) ⊗ SU(2). A thorough description of this noncompact *spectrum-generating* dynamical *Lie group* appeared in the preceding chapters.

At this point, it suffices to state that the symmetry group $\mathcal{G}$ consists of *operators* $\hat{A}$, $\hat{B}$, $\hat{C}$, and so on, that act on the states of $\mathscr{S}$ to *transform* them into one another. Worded somewhat more concretely, the group $\mathcal{G}$ is said to be capable of *transmuting* every element into every other. It could, on this account be compared with the *philosophers' stone* (or *lapis philosophorum*) of the Medieval alchemists—a legendary substance that had the ability to turn base metals (e.g., lead) into noble metals (e.g., gold or silver). It should be emphasized however, that this analogy is merely used as a convenient *metaphor*; the transformations, induced by $\mathcal{G}$, are abstract mathematical transformations that have no real counterpart in our physical world.

Returning, for a moment, to the hylomorphic doctrine of Aristotle, the symmetry group $\mathcal{G}$ possesses the power to *inform* the primary matter with any form to create whatever element. It is the *common cause* of all elemental transformations by *annihilating* one particular form (i.e., reducing the element to the proto hyle) and *creating* another.

The periodic system

The periodic system results from a systematic *breaking* of the symmetry group $\mathcal{G}$ into smaller subgroups $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$, and so on. As will be expounded in the following chapter, the periodic system is described by the decreasing chain of subgroups:

$$\mathrm{SO}(4,2) \otimes \mathrm{SU}(2) \supset \mathrm{SO}(4,2) \supset \mathrm{SO}(3,2) \supset \mathrm{SO}(4)' \supset \mathrm{SO}(3). \qquad (13.28)$$

With each subsequent symmetry-breaking act, the states of the system $\mathscr{S}$ break down into *multiplets* and *submultiplets* until a complete classification of the states has been obtained.

From phase space to Hilbert space

For a *classical system*, such as the triangle in §1.2.2, the states of the system $\mathscr{S}$ could be represented by *points* in *phase space*. The superparticle $\mathscr{S}$, in contrast, will be represented in a Hilbert space, and the states α, β, γ, and so on, are considered to be *vectors* in this linear *vector space* $\mathscr{R}$ (see Table 13.6).

Table 13.6 Classical versus quantum space.

System ($\mathscr{S}$)	↔	*State representation*
Classical	↔	Points in phase space
Quantum	↔	Vectors in vector space

To distinguish states in phase space from states in Hilbert space let us adopt Paul Dirac's *bra-ket notation* and represent every quantum state by a *ket vector* $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, and so on in the vector space $\mathscr{R}$. The vector space $\mathscr{R}$ is said to be:

1. *Linear*: If $|\alpha\rangle$ and $|\beta\rangle$ are vectors in $\mathscr{R}$, then $|\alpha\rangle + |\beta\rangle$ also belongs to $\mathscr{R}$.
2. *Complex*: Let c be a complex number; then, $c|\alpha\rangle$ is also a vector of $\mathscr{R}$.[77]
3. *Infinite-dimensional*: The different chemical elements form a *basis* for the vector space $\mathscr{R}$. Since the number of elements is denumerably infinite (at least theoretically), $\mathscr{R}$ will be infinite-dimensional.

The probability of a *transition* from $|\alpha\rangle$ to $|\beta\rangle$ can be determined by defining the *inner product* of two states, denoted $\langle\alpha|\beta\rangle$.[78] This turns the vector space $\mathscr{R}$ into a *Hilbert space* $\mathcal{H}$. Then, if both $|\alpha\rangle$ and $|\beta\rangle$ are normalized, the *probability* of a transition from $|\alpha\rangle$ to $|\beta\rangle$ is given by the squared probability amplitude $|\langle\alpha|\beta\rangle|^2$.

Transitions are still effected under the action of operators $\hat{T}_A$, $\hat{T}_B$, $\hat{T}_C$, and so on, that form a *representation* $\{T_g\}$ of the group $\mathcal{G}$ called the *symmetry group* of the system $\mathscr{S}$. In this monograph, $\mathcal{G}$ will always be a *Lie group*. The representation $\{T_g\}$ is operative in the space $\mathscr{R}$ of the quantum system and is defined by the action of the operators $\hat{T}_A$, $\hat{T}_B$, and so on, on the basis vectors $|\alpha\rangle$, $|\beta\rangle$, and so on. That is, the basis vectors are said to form a basis for an infinite-dimensional irreducible representation (or *unirrep*) of the dynamical symmetry group $\mathcal{G}$.

All *quantum transitions* are carried out under the action of operators of the form $\hat{T}_A + i\hat{T}_B$, where $\hat{T}_A$ and $\hat{T}_B$ are *generators* of the *Lie algebra* $\mathfrak{g}$ (corresponding to the Lie group $\mathcal{G}$), and where the operators $\hat{T}_A + i\hat{T}_B$ constitute *ladder operators* from the *complex hull* of this Lie algebra.[79] As an example of such ladder operators, the *raising* and *lowering operators* $\hat{L}_x \pm i\hat{L}_y$ of the *angular momentum algebra* $\mathfrak{so}(3)$, which shift a state $|nlm\rangle$ to the new state $|nl(m\pm 1)\rangle$, could be mentioned.

Justification of the EPA

Earlier, the question was raised whether it is justified to ignore the atomic substructure completely. We argue in favor of the aforementioned approach inasmuch as it parallels the course taken in elementary particle physics during the 1960s. Neither Murray Gell-Mann nor Yuval Ne'eman (or any other physicist for that matter) had any clue regarding the possible *substructure* of the hadrons. Yet, this did not prevent them from proposing a group theoretical classification of the particle zoo (§13.7.1).

---

[77] Technically, $|\alpha\rangle$ and $c|\alpha\rangle$ (with c a complex number) represent the same quantum state. It is therefore better to speak of *rays* $c|\alpha\rangle$ rather than vectors $|\alpha\rangle$ as representing quantum states.

[78] The left part $\langle\alpha|$ is called a *bra vector*, and the inner product is called a *bracket*.

[79] The complex hull of a Lie algebra $\mathfrak{g}$ is obtained by taking the tensor product of $\mathfrak{g}$ with the complex field $\mathbb{C}$. This is also referred to as the *complexification* of $\mathfrak{g}$.

Admittedly, the eightfold way did point to a hadronic substructure that later led to the *quark hypothesis*. In much the same way, it is desirable to interpret our group theoretical analysis as pointing to the underlying quantum mechanics of many-electron systems.

But, "what is the *gain* of ignoring the detailed description of [the] atom? [emphasis added]," inquired Byakov and colleagues.[80] The EPA offers two important advantages:

1. First of all, the entire set of chemical elements is treated as a whole; they are considered to be states of a *single* quantum system, whereas quantum mechanics treats each element as a *separate* quantum system. In exactly the same sense did Gell-Mann's eightfold way consider families of baryons or mesons as states of a single quantum system.
2. Second, because there is no talk about quantum numbers, this approach might circumvent the problems encountered in the quantum mechanical analysis of the periodic system (see, in particular, §13.5).

13.6.3 System and system states

In the group theoretical approach adopted in this chapter, the different chemical elements are treated as *structureless* particles; atoms are assumed to be *noncomposite* and their internal dynamics can thus be ignored. Instead, the chemical elements are considered to be various *states* of a single *quantum system* $\mathscr{S}$, which in turn is treated as a kind of *superparticle*. Following a proposition made by Wulfman in 1978, this pseudoparticle, with a spectrum that is Barut's atomic supermultiplet, is denoted by the name *baruton*.[81] The different states (or elements) are represented by ket vectors $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, and so on, that form a *basis* for an infinite-dimensional Hilbert space $\mathcal{H}$.

Within the EPA, a set of group theoretical labels n, l, m_l, and m_s are then introduced to identify each of the different ket vectors, denoted $|nlm_lm_s\rangle$. These *internal quantum numbers* are not to be confounded with the one-electron quantum numbers of hydrogenic systems, unless additional hypotheses are introduced.[82] Rather, the labels n, l, m_l, and m_s find their origin in the postulation of a particular symmetry group $\mathcal{G}$ of the periodic table. Since the range of variation of these formal EPA labels is the same as for the physical APA quantum numbers of the hydrogen atom, the conformal SO(4,2) group serves as a natural candidate for our study of the global group structure of the periodic system.

In order not to be prejudiced about the hydrogenic quantum numbers, Barut introduced the symbols ν, λ, μ_λ, and μ_σ.[83] Ostrovsky similarly distinguished between the conventional quantum numbers and the abstract SO(4,2) labels by denoting the latter ones with a tilde: $\tilde{n}$, $\tilde{l}$, $\widetilde{m}_l$, and $\widetilde{m}_s$.[84] Although this symbolism breaks the link with quantum mechanics explicitly, we will not adopt this notation, but use the traditional n, l, m_l, and m_s labels instead.

---

[80] V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet, "Group-Theoretical Classification of Chemical Elements. I. Physical Foundations," p. 4.

[81] Wulfman, "Dynamical Groups in Atomic and Molecular Physics."

[82] Ostrovsky, "Group Theory Applied to the Periodic Table of the Elements," p. 269.

[83] Barut, "Group Structure of the Periodic System."

[84] Ostrovsky, "Group Theory and Periodic System of Elements," p. 196.

It should be emphasized, however, that even without this quantum mechanical underpinning, the numbers n, l, m_l, and m_s serve as natural labels for the elements. Their value can be inferred from the *position* of an element in the periodic table, and from this point of view, the ket $|nlm_lm_s\rangle$ serves as an *address* for the different elements.

13.6.4 Symmetry and symmetry breaking

As indicated previously, the overall symmetry group $\mathcal{G}$ of the periodic table will be identified with the direct product group SO(4,2) $\otimes$ SU(2). The special real pseudo-orthogonal group in $4+2$ dimensions (or conformal group) SO(4,2) was introduced in Chapter 12 as a semisimple Lie group of order $r = 15$ and rank $l = 3$. This group provides the labels n, l, and m_l. To introduce the fourth variable, m_s, the direct product is formed between SO(4,2) and the special unitary group in two dimensions, SU(2). This semisimple Lie group of order $r = 3$ and rank $l = 1$ then serves to label the spin state of an element. The introduction of this additional "spin" group thus leads to a doubling of the dimensions of the different unirreps of the SO(4,2) subgroups.

The chemical elements form a basis for an infinite-dimensional unitary irreducible representation denoted $h \otimes [2]$, of the SO(4,2) $\otimes$ SU(2) group, where h is the unirrep of SO(4,2) and [2] stands for the fundamental representation of the spectral SU(2) group. The entire set of chemical elements is, in this way, housed in a single (infinite-dimensional) manifold of the symmetry group. From the point of view of the $\mathrm{SO}(4,2) \otimes \mathrm{SU}(2)$ group, the chemical elements are thus perfectly equivalent and interrelated.

To see this, raising and lowering operators can be constructed from the generators of the corresponding spectrum-generating Lie algebra $\mathfrak{so}(4,2) \oplus \mathfrak{su}(2)$. The resulting shift operators are able to transform any state $|nlm_lm_s\rangle$ into any other state $\left|n'l'm_l'm_s'\right\rangle$. When written in the Cartan-Weyl basis, the action of these ladder operators can be depicted in a *root diagram*. This was pursued in Chapter 12 and resulted in a cuboctahedral arrangement of the Weyl generators. To put it more poetically, the $\mathfrak{so}(4,2) \oplus \mathfrak{su}(2)$ root diagram represents the *philosophers' stone of quantum alchemy*; it enables one to *transmute* every element into every other.

Shattered symmetry

Only when the philosophers' stone is shattered—that is to say, when the group $\mathcal{G} =$ SO(4,2) $\otimes$ SU(2) is broken into smaller subgroups $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$, and so on—does the infinite-dimensional manifold (i.e., the baruton) break up into smaller multiplets.

In this sense, the baruton can be said to represent the *primeval atom*, at the point of the Big Bang, when all energies were degenerate. As its symmetry broke, the universe unfolded and phenomena appeared. Following Pierre Curie's dictum, it is precisely *because* the symmetry is broken that we are able to distinguish the phenomena. The observable manifestations of the baruton are the different chemical elements, which are arranged together in the periodic system. This implies that the structure of the periodic table is to be found in a particular *symmetry breaking* of the $\mathrm{SO}(4,2) \otimes \mathrm{SU}(2)$ group, which involves the identification of the different subalgebras of $\mathfrak{so}(4,2)$. The degeneracy pattern of the periodic system thus follows from the decomposition of the global multiplet over the subgroup unirreps.

13.6.5 Mass formulas

In the APA to the hydrogen atom, formulas could be derived for the system's energy. With the EPA, this is no longer possible and additional formulas have to be introduced to describe the *energy-level ordering*. These formulas find their origin outside the group theoretical scheme. A distinction can be made between *ordering formulas*, which yield the relative order of the energy levels as a function of their group theoretical labels, and *mass formulas*, which quantify the energy (or mass) of each degeneracy level.

The resulting *mass operator* $\hat{M}$ (known variously as a *symmetry-breaking operator*) belongs to the *universal enveloping algebra* $U(\mathfrak{g})$ of the system's symmetry group $\mathcal{G}$; it is typically expressed as a polynomial in the Cartan generators $\hat{H}_i$ of $\mathfrak{g}$ (written in its Cartan-Weyl basis) and Casimir operators $\hat{C}_\mu$ of the subgroups $\mathcal{G}_1$, $\mathcal{G}_2$, and so on. The eigenkets of $\hat{M}$ are all the vectors $|nlm_lm_s\rangle$; their corresponding eigenvalues are the positive integer numbers 1, 2, 3, and so on of multiplicity 1.[85] This implies a *one-to-one correspondence* between the set of quantum numbers n, l, m_l, m_s and the eigenvalues of $\hat{M}$, which are identified with the atomic number Z. Each chemical element with atomic number Z is thus ascribed an *address* in terms of the group theoretical labels of the symmetry group $\mathcal{G}$:

$$Z \leftrightarrow \{n, l, m_l, m_s\}. \tag{13.29}$$

Based on this association, each chemical element can be accommodated in a group theoretical version of the periodic system, which can then be compared with the conventional Mendeleev chart. This procedure is similar to the approach taken in hadron physics, in which mass operators have been analogously defined for the SU(3) and SU(6) groups (cf. *Gell-Mann–Okubo mass formula*). It should be emphasized once more that the obtained mass formulas are purely *phenomenological*; they are often superseded as theoretical research in quantum mechanics (or quantum chromodynamics) advances.

It should also be noted that mass formulas can be cast as *model Hamiltonians* expressed in terms of the generators of the spectrum-generating algebra.[86] The construction of such *formal* Hamiltonians differs from the *physical* Hamiltonians in the APA.

13.7 LITERATURE STUDY

Several authors have attempted to provide a group theoretical articulation of the periodic system of chemical elements. Various *groups* and *chains of subgroups* have been proposed to this aim, and they are enumerated in Table 13.7.

13.7.1 Historical prelude

The mathematical theory of abstract groups had been used as a *classificatory tool* in the 1960s by Murray Gell-mann and Yuval Ne'eman to classify the zoo of elementary particles. This group theoretical approach led to the *eightfold way* and the discovery

---

[85] A. I. Fet. "Conformal Symmetry of the Chemical Elements." *Theoretical and Mathematical Physics* 22.3 (1975), pp. 227–235.

[86] Ostrovsky, "Group Theory Applied to the Periodic Table of the Elements," p. 269.

Table 13.7 Proposed symmetry groups of the periodic system, along with their chain of subgroups. The different entries have been grouped according to the four schools introduced in the literature study. Later developments by Kibler and Thyssen and colleagues[a] are also included.

Author	*Year*	*Chain of subgroups*
Barut	1972	$SO(4,2) \supset SO(3,2) \supset SO(3) \otimes O(2)$
Odabaşi	1973, 2010	$SO(4,2) \supset SO(3,2) \supset SO(3) \otimes O(2)$
Ostrovsky	1980, 1981, 2006	$O(4,2) \otimes SU(2)_S \otimes SU(2)_T \supset O(4,2) \otimes SU(2)_S \supset O(4) \otimes SU(2)_S \supset O(3)$
Novaro and Berrondo	1972	$SU(2) \otimes SU(2) \otimes SU(2) \supset O(4) \supset SO(3)$
Berrondo and Novaro	1973, 1989, 2006, 2010	$E(4) \supset SU(2) \otimes SU(2) \otimes SU(2) \supset O(4) \supset SO(3)$
Rumer and Fet	1971	$Spin(4) \supset SU(2)_M, SU(2)_C$
Konopel'chenko	1972	$SO(2,4) + R \supset SO(2,4) \supset SO(4) \supset SO(3) \supset SO(2)$
Fet	1974, 1975	$SO(4,2) \otimes SU(2) \supset Spin(4) \otimes SU(2) \supset SU(2)_C \otimes SU(2)$
Byakov et al.	1976, 1977	$SO(4,2) \otimes SU(2) \supset Spin(4) \otimes SU(2) \supset SU(2)_C \otimes SU(2)$
Fet	1979, 1980, 1981	$O(4,2) \otimes SU(2) \otimes Z_2$
Fet	1989, 1992, 1996, 2010	$O(4,2) \otimes SU(2) \otimes SU(2)$
Kibler and Negadi	1989, 2004, 2006, 2007	$SO(4,2) \otimes SU(2) \supset SO(4) \otimes SU(2) \supset SO(3) \otimes SU(2) \supset SU(2)$
Thyssen and Ceulemans	2013	$SO(4,2) \otimes SU(2) \supset SO(3,2) \otimes SU(2) \supset SO'(4) \otimes SU(2)$

[a] P. Thyssen. "Symmetry and Symmetry Breaking in the Periodic Table: Towards a Group-Theoretical Classification of the Chemical Elements." PhD diss., Katholieke Universiteit Leuven, 2013.

of the quark structure of hadrons, both of which were described by the unitary SU(3) group.[87]

Following the particle physics tradition, a handful of scientists deemed it probable that there existed a more sophisticated, *symmetry-based* way of understanding how the chemical elements should be accommodated in the periodic table. Some even believed that group theory could shed some light on the way the periodic law emerges from its quantum mechanical foundations.

The phenomenological study of the global group structure of the periodic system thus originated during the 1970s with the pioneering work of a small group of theoretical physicists. As so often happens in the history of science, *multiple independent discoveries* typically occur when the time is ripe.[88] The application of group theory to the periodic system was no exception; at least four independent research groups in Novosibirsk (USSR), Boulder (Colorado), Mexico City (Mexico), and St. Petersburg (USSR) simultaneously saw the potential uses of group theory to explain the structure of Mendeleev's table.

A very first (but rather vague) suggestion in this direction came, however, from German physicist D. Neubert, who believed that "the high symmetry of the [periodic system could] be of interest from different point of views."[89]

Rumer, Fet, and Konopel'chenko (Novosibirsk, USSR)

In the same year, the renowned Soviet physicist Yurij Borisovich Rumer (1901–1985) and mathematician Abram Ilyich Fet (1924–2007) published a book on *The Theory of Unitary Symmetry Groups*.[90] Rumer had been an assistant of Max Born (1882–1970) during his internship at the University of Göttingen from 1929–1932. After serving as a professor at the Moscow State University from 1932 untill 1938, Rumer was arrested as an accomplice of Lev Landau (1908–1968) and was sent to Soviet camps. More than a decade later, Rumer finally moved to Novosibirsk, where he intended to apply the principles of group theory and representation theory to problems in biology and chemistry.[91]

In the field of biology, Rumer studied the symmetries of the *genetic code*, along with B. G. Konopel'chenko. Around the same time, Rumer wrote his book with Fet on unitary symmetry groups, dedicated to the symmetries of elementary particle physics. The writing of this monograph stimulated Rumer and Fet to embark on a "non-traditional [project]—the group theoretical description of the system of

---

[87] Gell-Mann and Ne'eman, *The Eightfold Way*.

[88] According to Robert K. Merton, these *multiples*, rather than *singletons* (i.e., a discovery made by a single individual), represent the common *pattern* in science. See R. K. Merton. "Singletons and Multiples in Scientific Discovery: A Chapter in the Sociology of Science." *Proceedings of the American Philosophical Society* 105.5 (1961), pp. 470–486. Or, to put it more poetically: "When the time is ripe for certain things, these things appear in different places in the manner of violets coming to light in early spring." As remarked by Farkas Bolyai, letter to his Janos, dated spring 1825.

[89] Neubert, "Double Shell Structure of the Periodic System of the Elements," p. 216.

[90] Yu B. Rumer and A. I. Fet. Teoriya Unitarnoi Simmetrii [*The Theory of Unitary Symmetry*]. Moscow: Nauka, 1970, p. 400. Another result of their fruitful collaboration was the book Yu B. Rumer and A. I. Fet. *Group Theory and Quantum Fields*. Moscow: Nauka, 1977, p. 248.

[91] A. I. Fet. *Symmetry Group of Chemical Elements*. Ed. R. G. Khlebopros. Novosibirsk: Nauka, 2010 [in Russian].

chemical elements."[92] Inspired by the group theoretical classification schemes in hadron physics, based on the unitary symmetries SU(3) and SU(6),[93] Rumer and Fet wanted to apply the same EPA principles to the manifold of chemical elements. Their first article appeared in 1971 in the journal *Theoretical and Mathematical Physics* and offered a detailed description of the periodic system based on a *two-sheeted covering* of the SO(4) group[94]:

$$\mathcal{G}_{RF} = \text{Spin}(4). \tag{13.30}$$

Notice that, with this choice, Rumer and Fet did not yet invoke noninvariance groups. The Spin(4) group was first introduced by Brauer and Weyl in 1935,[95] and it corresponds to the group of orthogonal transformations (with det = 1) that leave the metric form $\xi_1^2 + \xi_2^2 + \xi_3^2 + \xi_4^2$ invariant in four-dimensional Euclidean space $\mathbb{R}^4$ (Chapter 10). Two different subgroups, denoted $\text{SU}(2)_\text{M}$ and $\text{SU}(2)_\text{C}$, were considered:

$$\text{Spin}(4) \supset \text{SU}(2)_\text{M}; \tag{13.31}$$

$$\text{Spin}(4) \supset \text{SU}(2)_\text{C},\ \text{SU}(2)'_\text{C}. \tag{13.32}$$

The first reduction gives rise to the *mechanical subgroup* $\text{SU}(2)_\text{M}$, which corresponds to the proper rotation subgroup SO(3) of SO(4), and is defined in the subspace $\mathbb{R}^3$ formed by the ξ_1, ξ_2, and ξ_3 coordinates. The corresponding $\mathfrak{su}(2)_\text{M}$ algebra is generated by the components of the angular momentum vector $\hat{\mathbf{L}}$.

The second reduction in Eq. (13.32) yields one of two equivalent *chemical subgroups*, denoted $\text{SU}(2)_\text{C}$ or $\text{SU}(2)_{\text{C}'}$, with Lie algebras that are generated, respectively, by the components of the $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ operator, with $\hat{J}_{1i} = 1/2\big(\hat{L}_i + \hat{A}_i\big)$ and $\hat{J}_{2i} = 1/2\big(\hat{L}_i - \hat{A}_i\big)$, as defined in Chapter 9. Both algebras appear in the direct sum decomposition of the $\mathfrak{so}(4)$ algebra: $\mathfrak{so}(4) = \mathfrak{su}(2)_\text{C} \oplus \mathfrak{su}(2)_{\text{C}'}$.

Rumer and Fet maintained that the hydrogenic states form a representation with respect to the first reduction, yielding the familiar quantum numbers n, l, and m_l. The chemical elements, on the other hand, were taken as basis states for an infinite-dimensional unitary representation of the covering group Spin(4) with respect to the second reduction. This representation was termed a *Coulomb system* for obvious reasons, and the obtained classification of the elements was claimed to be in accordance with their natural grouping in the periodic system.

This description, however, suffers from the following shortcomings:

1. First, the representations are *reducible*. Because the chemical elements are grouped in different pairwise, disjoint multiplets of the Spin(4) group, the *Coulomb system* cannot be regarded as *elementary* in the group theoretical sense.[96]

---

[92] Quoted from a letter Rumer wrote to academician M. A. Leontovich. Ibid., p. 4.

[93] This was described by the authors in Rumer and Fet, *Teoriya Unitarnoi Simmetrii* [*The Theory of Unitary Symmetry*].

[94] Rumer and Fet, "The Group Spin (4) and the Mendeleev System."

[95] R. Brauer and H. Weyl. "Spinors in *n* Dimensions." *American Journal of Mathematics* 57.2 (1935), pp. 425–449.

[96] Fet, "Conformal Symmetry of the Chemical Elements," p. 227.

2. Second, both subalgebras, $\mathfrak{su}(2)_{\mathrm{M}}$ and $\mathfrak{su}(2)_{\mathrm{C}}$, are *algebraically identical*, as indicated by Ostrovsky.[97] Within the EPA, these subalgebras are indistinguishable, and it is not clear in which sense $\mathfrak{su}(2)_{\mathrm{C}}$ could be related to the periodic system, and $\mathfrak{su}(2)_{\mathrm{M}}$ to the hydrogen atom, as long as their generators are not realized in terms of physical operators, as indicated earlier.
3. Finally, and most important, albeit a one-to-one correspondence is established between the group labels and the elements, no mass formula is provided, and neither does the proposed description rationalize (or even mention) the doubling of the periods or the degeneracy pattern resulting from the Madelung $(n+l,n)$ rule.

These initial ideas were further developed by Fet during the following ten years in a number of papers and conference proceedings, which greatly improved his description of the periodic system. The SO(4) group was broadened to the *conformal* SO(4,2) group by Konopel'chenko in 1972,[98] and in 1975, Fet proposed the direct product group SO(4,2) ⊗ SU(2).[99] To be correct, Fet proposed the group

$$\mathcal{G}_{\mathrm{F}} = \mathrm{S}\widetilde{\mathrm{O}}(4,2) \otimes \mathrm{SU}(2), \tag{13.33}$$

with $\mathrm{S}\widetilde{\mathrm{O}}(4,2)$ the *universal covering* of the conformal group.[100] The introduction of this noninvariance group eliminated the first shortcoming mentioned previously. This

---

[97] Ostrovsky, "Group Theory Applied to the Periodic Table of the Elements," p. 278.

[98] B. G. Konopel'chenko. "Gruppa SO(2,4)+R i Tablitza Mendeleeva [The SO(2,4)+R Group and Mendeleev's Table]." *Novosibirsk*: Preprint of the Institute of Nuclear Physics IYaF-40.72, 1972 [in Russian]. See also B. G. Konopel'chenko and Yu B. Rumer. "Atomy i Adrony (Problemy Klassifikacii) [Atoms and Hadrons (Classification Problems)]." *Soviet Physics Uspekhi* 22.10 (1979), pp. 837–840. Translated from the Russian: B. G. Konopel'chenko and Yu B. Rumer. "Atoms and Hadrons (Classification Problems)." *Uspekhi Fizicheskikh Nauk* 129 (1979), pp. 339–342.

[99] Fet, "Conformal Symmetry of the Chemical Elements." A popular account of Fet's work appeared in V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet. "Group-Theoretical Classification of Chemical Elements. I. Physical Foundations."; V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet. "Group-Theoretical Classification of Chemical Elements: II. Description of Applied Groups." Moscow, 1976; V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet. "Group-Theoretical Classification of Chemical Elements. III. Comparison with the Properties of Elements." Moscow, 1977. Fet laid special emphasis on A. I. Fet. *The System of the Elements from the Group-Theoretic Viewpoint.* 1979. Novosibirsk: Preprint of the Institute of Nuclear Physics IYaF-40.72. This is a preprint of the Institute of Inorganic Chemistry, Siberian Branch of the USSR Academy of Sciences, 1979, No 1. This preprint was later reproduced in A. I. Fet. "The System of Elements from the Group-Theoretic Viewpoint." In: *Periodic Systems and Their Relations to the Systematic Analysis of Molecular Data.* Ed. Ray Hefferlin. Lewiston, NY: The Edwin Mellen Press, 1989, pp. 41–86.

[100] A. I. Fet. "Konformnaja Gruppa i Himicheskoe Srodstovo [Conformal Group and Chemical Affinity]," *JETP Letters* 20.1 (1974), pp. 10–11. Translated from the Russian: A. I. Fet. "Conformal Group and Chemical Affinity." *Pis'ma v ZhETF* 20.1 (1974), pp. 24–26. See also Fet, "Konformnaja Simmetrija Himicheskih Jelementov [Conformal Symmetry of the Chemical Elements]," translated from the Russian: Fet, "Conformal Symmetry of the Chemical Elements." The same group description also appeared in three semipopular preprints ITEP-26, ITEP-90, and ITEP-7 of the *Moscow Institute of Theoretical and Experimental Physics.* See V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet., "Group-Theoretical Classification of Chemical Elements: I. Physical Foundations;" V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet. "Group-Theoretical Classification of Chemical Elements: II. Description of Applied Groups;" Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet. "Group-Theoretical Classification of Chemical Elements: III. Comparison with the Properties of Elements."

group was subsequently reduced over the following chain of subgroups:

$$\widetilde{\mathrm{SO}}(4,2) \otimes \mathrm{SU}(2) \supset \mathrm{Spin}(4) \otimes \mathrm{SU}(2) \supset \mathrm{SU}(2)_{\mathrm{C}} \otimes \mathrm{SU}(2). \tag{13.34}$$

At this point, Fet had also become acquainted with Barut's paper,[101] which was likewise based on the SO(4,2) group and which explicitly mentioned the Madelung rule and period doubling. When Fet introduced the mass operator $\hat{A}$, he thus referred to the *lexicographical* ordering of the quantum numbers according to the mass formula

$$A = \frac{1}{6} d\left(d^2-1\right) + \frac{1}{2}(d+1)^2 - \frac{1}{2}\kappa(d)\cdot(d+1) - 2\left(\lambda^2+1\right) + 2\mu + s_3 + \frac{3}{2}, \tag{13.35}$$

where $d = n+\lambda$, and $\kappa(d)$ is 0 for odd d and 1 for even d. The quantum numbers n, λ, μ, and s_3 are defined similarly to n, l, m_l, and m_s.

The periodic system, proposed by Fet, is nevertheless hardly satisfying. The grouping based on the reduction in Eq. (13.34) yields SO(4)-like multiplets of dimension n^2, which are labeled by n rather than $n+\lambda$. The chemical elements are consequently forced into *artificial periods* of constant n, which are laid out in a horizontal way, with λ varying in the vertical direction. As a result, the Madelung order is not read from top to bottom and left to right, but in a *slanted* way, which is far from transparent. The introduction of an (n,λ) framework, although natural for a classification of the hydrogenic states, is utterly inadequate for the construction of a periodic table, which necessitates an $(n+\lambda,\lambda)$ skeleton instead.

What is more, in using a classification founded on the quantum numbers n and λ (rather than $n+\lambda$ and λ), the spectral SU(2) group in Eq. (13.33) is merely invoked to double the number of states within each multiplet (comparable with the doubling of energy levels in the hydrogen spectrum resulting from the electronic spin). Although Fet referred to this as being a result of the *chemical spin* s_3, it clearly fails to account for the doubling of the periods (see also §13.4.2).[102] The later developments by Fet will be examined in §13.7.2.

In 1984, Fet wrote a complete monograph on the *Symmetry Group of Chemical Elements*.[103] After the entire book had been linotyped by the Siberian division of the Nauka publishing house, it was suddenly withdrawn from printing by an order of some academic functionary.[104] According to the author's widow, this order was solely political and had no relation to science. R. G. Khlebopros believed the reasons were to be found in Fet's personality:

> A talented mathematician and physicist, a very well-educated and intelligent person with a sense of dignity and independence, he was of course envied and hated by ungifted science bureaucrats.[105]

Two years later, Fet was fired from the institute "for lack of publications." In 1992, the book came out in abridged form in the collected works of *Mathematical Modelling*

---

[101] Barut, "Group Structure of the Periodic System."

[102] Fet seems to have been on the right track, though, because he claimed that s_3 "must be related to the well-known difference between the chemical properties of even and odd elements." Quoted from Fet, "Conformal Symmetry of the Chemical Elements," p. 230.

[103] As far as we can judge, this is the only book ever written on the group theory of the periodic system.

[104] Private communication with Ludmila Petrova-Fet, the author's widow. See also Fet, *Symmetry Group of Chemical Elements*, p. 6.

[105] Ibid., p. 6.

in Biology and Chemistry.[106] Only after Fet's death, and thanks to the efforts of Prof. Khlebopros, was the book finally published in its entirety as it was supposed to be in 1984.[107]

Barut (Boulder, Colorado, USA)

Undoubtedly Asim O. Barut (1926–1994) from the University of Colorado played a major role in the development of noninvariance groups, or "spectrum-generating groups," as he called them. As mentioned in Chapter 12, his visit to New Zealand in 1971 as an Erskine Fellow at the University of Canterbury, in Christchurch, New Zealand, yielded two important documents that explain the construction of SO(4,2) as the noninvariance group of hydrogen, and discuss its application to the Aufbau principle of the periodic table. The SO(4,2) construction is presented in a booklet containing notes of the lectures Barut gave as an Erskine Fellow.[108] In his contribution to the proceedings of the Rutherford symposium,[109] Barut considered symmetry breakings of SO(4,2), associated with the Aufbau principle. He noted the importance of the symmetry breaking of SO(4,2) to its subgroup SO(3,2). On reducing the SO(4,2) parent group to its subgroup SO(3,2), the infinite representation h, which contains all $|nlm_l\rangle$ states, splits into exactly *two* representations of SO(3,2). The state vectors $|nlm\rangle$ (i.e., chemical elements) with $n+l$ even (respectively, odd) span an infinite-dimensional *unirrep* of SO(3,2), denoted h_e (respectively, h_o).[110] This yields the following branching rule:

$$h = h_e \oplus h_o. \tag{13.36}$$

This explanation goes back to an earlier study by Barut and Bohm, in which the different representation classes of the SO(4,2) group were studied, along with their reduction over the chain[111]:

$$\mathrm{SO}(4,2) \supset \mathrm{SO}(3,2) \supset \mathrm{SO}(3) \otimes \mathrm{SO}(2). \tag{13.37}$$

It offers an attractive explanation for the doubling of the periodic table. Nonetheless, this approach has been criticized by Ostrovsky, who noted that "according to this reduction the O(4) subgroup completely loses its significance."[112] Ostrovsky continued: "the label n and the [unirrep] dimension n^2 are ultimately related to the

---

[106] A. I. Fet. "Gruppa Simmetrii Khimicheskikh Elementov [Symmetry Group of Chemical Elements]." *Matematicheskoe Modelirovanie v Biologii i Khimii* [*Mathematical Modelling in Biology and Chemistry*]. Ed. R. G. Khlebopros. Novosibirsk: Nauka Publishers, 1992, pp. 118–203 [in Russian].

[107] Fet, *Symmetry Group of Chemical Elements* [in Russian].

[108] A. O. Barut. *Dynamical Groups and Generalized Symmetries in Quantum Theory* (*With Applications in Atomic and Particle Physics*). Ed. A. N. Brooks. University of Canterbury, Christchurch, New Zealand: Bascands Ltd., 1972. Barut published his first paper on SO(4,2) together with Hagen Kleinert back in 1967. A. O. Barut and H. Kleinert. "Transition Probabilities of the Hydrogen Atom from Non-compact Dynamical Groups. *Physical Review* 156(1967) 1541–1545.

[109] Barut, "Group Structure of the Periodic System."

[110] Here, we have adopted the notation introduced by Kibler in Kibler. "On the Use of the Group SO(4,2) in Atomic and Molecular Physics."

[111] See Barut. "Group Structure of the Periodic System." and A. O. Barut and A. Böhm. "Reduction of a Class of O(4,2) Representation with Respect to SO(4,1) and SO(3,2)." *Journal of Mathematical Physics* 11.10 (1970), pp. 2938–2945.

[112] Ostrovsky. "Group Theory Applied to the Periodic Table of the Elements," p. 277.

O(4) group. They are critically important for the description of the periodic table; and yet they *do not appear* in Barut's scheme" [emphasis in original].[113]

In defense of Barut's approach, it should be emphasized that it is not the principal quantum number n, but the Madelung quantum number $N = n + l$ that is of importance to the periodic system. Admittedly, the different $n + l$ multiplets exhibit the same hydrogenic dimensions (2, 8, 18, 32, …), but they cannot be related to a straightforward SO(4) group, as will be explained in Chapter 14.

On a more positive note, Novaro referred to Barut's approach as "the most successful group-theoretical explanation of the Aufbau scheme."[114] He considered the chain of groups in Eq. (13.37) promising enough to provide it with a geometric interpretation in terms of an axially symmetric top in three dimensions.[115] This, however, did not withhold Novaro from submitting Barut's results to a critical examination. It is interesting to note that Novaro rewrote Eq. (13.37) as

$$\mathrm{SO}(4,2) \supset \mathrm{SO}(3,2) \supset \mathrm{O}(4), \tag{13.38}$$

claiming that "Barut's chain of subgroups does contain the symmetry group O(4)."[116] As Ostrovsky noted, this equation cannot be correct because the SO(3,2) group does *not* contain SO(4) as a subgroup. The $\mathfrak{so}(4)$ algebra is generated by the $\hat{\mathbf{L}}$ and $\hat{\mathbf{A}}$ operators, but the components of $\hat{\mathbf{A}}$ are lost in the reduction of $\mathfrak{so}(4,2)$ over the $\mathfrak{so}(3,2)$ Lie algebra.

Impressed by Barut's group theoretical treatment of the periodic system, Halis Odabaşi (1931–2011) felt inclined to explore these ideas somewhat further. At the Sanibel Symposium of 1973,[117] which was held in honor of Edward Uhler Condon (1902–1974), Odabaşi presented convincing evidence for the chain of subgroups that had been proposed by Barut back in 1971.[118]

Novaro, Wolf, and Berrondo (Mexico City)

The Mexican scientist Octavio Novaro, who was also present at the Sanibel Symposium, picked up the thread where Odabaşi had left off and provided "geometrical images" for Barut's chain of groups.[119] Novaro had obtained his doctoral degree in physics under the supervision of Marcos Moshinsky (1921–2009) in 1969. His first attempts at a group theoretical justification of Bohr's Aufbau scheme started in 1971, when he convinced his office colleague Bernardo Wolf to team up with him.[120]

---

[113] ibid., 277.

[114] O. Novaro. "Group Theoretical Aspects of the Periodic Table of the Elements." *Journal of Molecular Structure: THEOCHEM* 199 (1989), p. 109.

[115] O. Novaro. "Comment on the Group Theoretical Justification of the Aufbau Scheme." *International Journal of Quantum Chemistry* 7.S7 (1973), pp. 53–56.

[116] Novaro, "Group Theoretical Aspects of the Periodic Table of the Elements," p. 109.

[117] The Sanibel Symposium was an international scientific conference on atomic, molecular, and solid-state physics and quantum biology. It was founded by Per-Olov Löwdin in 1960 and was held every winter on Sanibel Island off the Gulf Coast of Florida.

[118] Odabaşi. "Some Evidence about the Dynamical Group SO (4,2): Symmetries of the Periodic Table of Elements." and Condon and Odabaşi, "The Order of Electron Shells for Atoms and Ions."

[119] Novaro, "Comment on the Group Theoretical Justification of the Aufbau Scheme."

[120] All historical recollections in this section are taken from O. Novaro. "Symmetries of the Periodic System." *AIP Conference Proceedings* 1323.1 (2010), pp. 244–256.

Starting from the groundbreaking work of Soviet physicist Vladimir A. Fock (1898–1974), who had explained the *accidental degeneracy* of the hydrogen atom in terms of a four-dimensional symmetry (Chapter 9), Novaro and Wolf subsequently tried to break this O(4) symmetry.[121] Despite the cold reaction from Moshinsky, Novaro's work attracted the attention of Per-Olov Löwdin who invited him to speak at the previously mentioned symposium.[122]

With the help of Carl Wulfman, Elpidio Chacón, and A. Freyre, Novaro forged ahead and elucidated in which way the O(4) and U(3) symmetries were broken in the first and second periods of Mendeleev's table.[123] However fascinating, this research only galvanized Novaro's curiosity as to whether there existed a *global symmetry* for the periodic system. Being familiar with Barut's work, Novaro felt it lacking in its description of the *period doubling*, as he considered the period doubling to be of type 2, 8, 8, 18, 18, 32, 32, and so on, instead of 2, 2, 8, 8, 18, 18, 32, 32, and so on.

Around that time, Manolo Berrondo returned to Mexico after obtaining a PhD with Löwdin. Soon after, Novaro and Berrondo embarked on their quest for the hidden symmetries of the periodic system, which they approached in a manner similar to the EPA discussed earlier.[124] They essentially considered the breaking of the O(4) symmetry of the Coulomb potential, which, we recall, is locally isomorphic to the product group $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$. A generalization of this led to the proposition of the direct product group

$$\mathcal{G}_{\mathrm{NB}} = \mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{SU}(2), \tag{13.39}$$

which resulted from three mutually commuting "angular momenta," denoted $\hat{\mathbf{P}}$, $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$, and obeying the following commutation rules:

$$\begin{gathered} \left[\hat{P}_i, \hat{P}_j\right] = i\varepsilon_{ijk}\hat{P}_k, \quad \left[\hat{Q}_i, \hat{Q}_j\right] = i\varepsilon_{ijk}\hat{Q}_k, \quad \left[\hat{R}_i, \hat{R}_j\right] = i\varepsilon_{ijk}\hat{R}_k, \\ \left[\hat{P}_i, \hat{Q}_j\right] = \left[\hat{Q}_i, \hat{R}_j\right] = \left[\hat{R}_i, \hat{P}_j\right] = 0, \quad \forall i, j = 1, 2, 3. \end{gathered} \tag{13.40}$$

This group was then reduced over the chain of groups:

$$\mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{SU}(2) \supset \mathrm{O}(4) \supset \mathrm{SO}(3). \tag{13.41}$$

---

[121] O. Novaro and K. B. Wolf. "A Model Hamiltonian for the Periodic Table." *Revista Mexicana de Fisica* 20 (1971), pp. 265–268.

[122] In his reminiscences of that period, Novaro later wondered whether Moshinsky's close friendship with Vladimir Fock might have explained his reservations. Perhaps he felt "we didn't do justice to [Fock's] elegant and esthetic theory," said Novaro, in Novaro. "Comment on the Group Theoretical Justication of the Aufbau Scheme," p. 244.

[123] O. Novaro. "Validity of O(4) Symmetry in 2nd Row Atoms." *Physics Letters A* 33.2 (1970), pp. 109–110; E. Chacon, M. Moshinsky, O. Novaro, and C. Wulfman. "O(4) and U(3) Symmetry Breaking in the 2s-2p Shell." *Physical Review A* 3.1 (1971), pp. 166–179; O. Novaro and A. Freyre. "O(4) and U(3) Symmetry Breaking in the Second Row of the Periodic Table." *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics* 20.5 (1971), pp. 861–871.

[124] O. Novaro and M. Berrondo. "Approximate Symmetry of the Periodic Table." *Journal of Physics B: Atomic and Molecular Physics* 5.6 (1972), pp. 1104–1110. These and other results have been reviewed in a number of other papers. See Novaro, "Group Theoretical Aspects of the Periodic Table of the Elements"; O. Novaro. "Group Theory of the Periodic Table." In: *The Mathematics of the Periodic Table.* Eds. Dennis H. Rouvray and R. Bruce King. New York: Nova Science Publishers, 2006, pp. 217–235; Novaro, "Symmetries of the Periodic System."

Table 13.8 Unitary irreducible representations of the group $\mathcal{G}_{\mathrm{NB}} = \mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ and its subgroups O(4) and SO(3). The dimensions of the $\mathcal{G}_{\mathrm{NB}}$ *unirreps* are listed in the last column.

unirrep $\mathcal{G}_{NB}$	$\supset$	*unirrep* *O(4)*	$\supset$	*unirrep* *SO(3)*	*Degeneracy*
$(0,0,0)$		$(0,0)$		$l=0$	1
$\left(\frac{1}{2},0,\frac{1}{2}\right)$		$\left(\frac{1}{2},\frac{1}{2}\right)$		$l=0,\ 1$	4
$\left(\frac{1}{2},\frac{1}{2},0\right)$		$(1,0)^+ \oplus (0,0)$		$l=0,\ 1$	4
$(1,0,1)$		$(1,1)$		$l=0,\ 1,\ 2$	9
$(1,1,0)$		$(2,0)^+ \oplus (1,0)^+ \oplus (0,0)$		$l=0,\ 1,\ 2$	9
$\left(\frac{3}{2},0,\frac{3}{2}\right)$		$\left(\frac{3}{2},\frac{3}{2}\right)$		$l=0,\ 1,\ 2,\ 3$	16

Adapted from O. Novaro and M. Berrondo. "Approximate Symmetry of the Periodic Table." *Journal of Physics B: Atomic and Molecular Physics* 5.6 (1972), pp. 1104–1110, p. 1108.

The *unirreps* of $\mathcal{G}_{\mathrm{NB}}$ are written as (p,q,r) with p, q, and r arising from the eigenvalues of the Casimir operators $\hat{P}^2$, $\hat{Q}^2$, and $\hat{R}^2$, and taking integer and semi-integer values. The *unirreps* of the O(4) subgroup can be introduced by *coupling* the first two angular momenta according to $\hat{\mathbf{P}} + \hat{\mathbf{Q}} = \hat{\mathbf{M}}'$. The resulting vector $\hat{\mathbf{M}}'$ is then coupled with the original $\hat{\mathbf{R}}$ to yield $\hat{\mathbf{L}} = \hat{\mathbf{M}}' + \hat{\mathbf{R}}$. The vectors $\hat{\mathbf{M}}$ and $\hat{\mathbf{R}}$ generate the O(4) group, and the components of $\hat{\mathbf{L}}$ form the SO(3) subgroup in Eq. (13.41).

To obtain the degeneracy structure of the periodic table, two *physical restrictions* have to be imposed. First, the vector $\hat{\mathbf{L}}$ must be interpreted as the total angular momentum of the system. This limits l to *integer* values. Second, each value of l is allowed to occur only once in each $\mathcal{G}_{\mathrm{NB}}$ multiplet. The number of admitted *unirreps* (p,q,r) is severely restricted under these conditions. The physically allowed *unirreps* are of one of two forms: $(p,p,0)$ or $(p,0,p)$.[125] A list of the first few *unirreps* is given in Table 13.8. Three disadvantages of this approach can be mentioned:

1. First, the chemical elements form a *reducible* representation under $\mathcal{G}_{\mathrm{NB}}$, as was the case with the Spin(4) group proposed by Rumer and Fet. Instead of being grouped in a single infinite-dimensional manifold of some noncompact noninvariance group, the elements belong to an infinitude of finite-dimensional multiplets of the compact $\mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ invariance group.
2. Second, the dimensions of the $\mathcal{G}_{\mathrm{NB}}$ *unirreps* in Table 13.8 are doubled, except for the very first multiplet, with a dimension 2 that occurs only once. The series 2, 8, 8, 18, 18, 32, 32 thus obtained corresponds to the cardinalities of the *chemical* periods in most conventional periodic tables.[126] In this book, however, the format of the left-step table is favored with period lengths 2, 2, 8, 8, 18, 18, 32, 32 that are doubled over the entire line. The doubling of the

---

[125] This is reminiscent of the *unirreps* (j,j) of the SO(4) group, which were obtained for the hydrogen atom under the restriction $j_1 = j_2$ (cf. Chapter 9).

[126] In this way, the atomic *magic numbers* can be derived, corresponding to the atomic numbers of the noble gases: He ($Z=2$), Ne ($Z=10$), Ar ($Z=18$), Kr ($Z=36$), Xe ($Z=54$), and Rn ($Z=86$).

first period cannot be explained by the group theoretical scheme proposed by Novaro and Berrondo.[127]

3. Finally, to obtain the previously mentioned dimensions, two *physical restrictions* had to be imposed that necessarily lie outside the EPA.

The first point was later remedied by the introduction of the Euclidean group in four dimensions, E(4), which contains $\mathcal{G}_{\text{NB}}$ as a subgroup.[128] With regard to the third point, Novaro is essentially facing the challenges posed in §13.6.5: How can one make the step from EPA to APA? It is as yet unclear how the proposed scheme relates to the internal dynamics of atomic physics. The quantum numbers obtained from the chain in Eq. (13.41), as well as their corresponding operators $\hat{\mathbf{P}}$, $\hat{\mathbf{Q}}$, $\hat{\mathbf{R}}$ lack a physical interpretation, and the dynamics of the symmetry-breaking mechanism are unknown, although the breaking is most probably a result of electron correlations and spin–orbit coupling effects. Novaro and Berrondo resolved this problem in part by constructing an *effective model Hamiltonian* in terms of the Casimir operators of the group $\mathcal{G}_{\text{NB}}$[129]:

$$\hat{\mathscr{H}}_{\text{eff}} = \frac{-\frac{1}{2}Z^2}{2\left[\left(\alpha\hat{\mathbf{P}} + \beta\hat{\mathbf{Q}}\right)^2 + \beta\hat{\mathbf{R}}^2\right] + 1}, \tag{13.42}$$

which corresponds to the Hamiltonian of a quantal symmetric top in 4-space.

Despite this joint effort, they did not settle the entire issue. In 1989, Novaro concluded that “despite many attempts and some advances, it is still necessary to accept that Löwdin’s call for a more sustained effort to understand the periodic system of the elements is still a valid challenge for theoretical chemists and physicists.”[130] Although Novaro had hoped to revive his efforts with Berrondo in 2009 to obtain a final answer to Löwdin’s plea, his hopes were not fulfilled.[131]

As has been observed before, Novaro provided a group theoretical interpretation of the cardinalities of the *chemical* periods, as given by the series of atomic *magic numbers* 2, 8, 8, 18, 18, 32, 32. “These magic numbers are not merely a duplication of hydrogen closed shells,” remarked Novaro, “[because] the number 2 only appears once, for helium.”[132] The doubling of the Aufbau scheme, as proposed by Barut, by invoking two different representations of the SO(3,2) group, was therefore considered by Novaro to be a “fatal flaw.”[133] The obtained dimensionalities do not “correspond to the magic numbers” and “as this chain of groups has been used implicitly by several authors, all [these] studies . . . necessarily lack this important aspect.”[134]

---

[127] Novaro and Berrondo were certainly aware of this fact, but they considered the conventional format of the periodic table to be more fundamental.

[128] M. Berrondo and O. Novaro. “On a Geometrical Realization of the Aufbau Scheme.” *Journal of Physics B: Atomic and Molecular Physics* 6.5 (1973), pp. 761–769.

[129] Novaro and Berrondo. “Approximate Symmetry of the Periodic Table,” p. 1109. and Berrondo and Novaro. “On a Geometrical Realization of the Aufbau Scheme.” See also Novaro and Wolf. “A Model Hamiltonian for the Periodic Table.”

[130] Novaro, “Group Theoretical Aspects of the Periodic Table of the Elements,” p. 117.

[131] Novaro, “Symmetries of the Periodic System,” p. 245.

[132] Novaro, “Group Theoretical Aspects of the Periodic Table of the Elements,” p. 104.

[133] Ibid., p. 110.

[134] Ibid., p. 112.

Kibler (Lyon, France)

Although Kibler cannot be counted among the pioneers who started the group theoretical study of the periodic system, he has been an ardent advocate of using the SO(4,2) ⊗ SU(2) group as a starting point for all group theoretical articulations of the periodic law.[135] Kibler has been instrumental in disseminating this research among a broader scientific audience. His articles typically have an educational flavour—introducing the many group theoretical concepts needed to tackle the periodic system.

Kibler heavily relied on the results obtained by the Russian school and thus proposed the symmetry group

$$\mathcal{G}_{\mathrm{K}} = \mathrm{SO}(4,2) \otimes \mathrm{SU}(2), \tag{13.43}$$

with the following chain of subgroups:

$$\mathrm{SO}(4,2) \otimes \mathrm{SU}(2) \supset \mathrm{SO}(4) \otimes \mathrm{SU}(2) \supset \mathrm{SO}(3) \otimes \mathrm{SU}(2) \supset \mathrm{SU}(2). \tag{13.44}$$

The periodic system offered by Kibler is identical to Fet's table, but has been rotated over 90°. It thus suffers the same shortcomings as Fet's suggestion. Ostrovsky likewise referred to the "highly irregular" filling and "inappropriate character" of Kibler's table.[136] He noted that "the shape of this chart does not reflect the $(n+l,n)$ rule with its characteristic period doubling …; the latter is actually lost at this stage in the construct."[137]

Kibler is probably one of the few researchers who insisted that group theory should not only be used in a qualitative sense to identify the overall structure of the periodic table, but also implies quantitative results. The associated Casimir operators indeed provide quantitative measures characterizing the states of the system, and these should be related to physical and chemical properties. This program was referred to as the KGR program because it was discussed at the Brewster's Kananaskis Guest Ranch in Canada during the 2003 Harry Wiener International Conference.[138]

---

[135] M. R. Kibler. "The Periodic System of Chemical Elements: Old and New Developments." *Journal of Molecular Structure: THEOCHEM* 187 (1989), pp. 83–93; M. R. Kibler. "On a Group-Theoretical Approach to the Periodic Table of Chemical Elements." *arXiv: quant-ph* 0408104. August 16, 2004. M. R. Kibler. "Classifying Chemical Elements and Particles: From the Atomic to the Sub-Atomic World." In: *The Periodic Table: Into the 21st Century*. Eds. D. H. Rouvray and R. B. King. Baldock: Research Studies Press, 2004, pp. 297–329; M. R. Kibler. "A Group-Theoretical Approach to the Periodic Table: Old and New Developments." *The Mathematics of the Periodic Table*. Eds. D. H. Rouvray and R. B. King. New York: Nova Science Publishers, 2006, pp. 237–263; M. R. Kibler. "From the Mendeleev Periodic Table to Particle Physics and Back to the Periodic Table." *Foundations of Chemistry* 9.3 (2007), pp. 221–234; M. R. Kibler. "Sur la Route de Mendeleïev: De la Chimie à la Physique des Particules." *La G@zette de l'IPNL* 14 (2007), pp. 7–8. See also M. R. Kibler and T. Negadi. "On the q-Analogue of the Hydrogen Atom." *Journal of Physics A: Mathematical and General* 24.22 (1991), pp. 5283–5289. T. Negadi and M. R. Kibler. "The Periodic Table in Flatland." *International Journal of Quantum Chemistry* 57.1 (1996), pp. 53–61; Kibler, "On the Use of the Group SO(4,2) in Atomic and Molecular Physics."

[136] Ostrovsky. "Group Theory Applied to the Periodic Table of the Elements," p. 280.

[137] Ibid. p. 279.

[138] Kibler, "A Group-Theoretical Approach to the Periodic Table of the Chemical Elements: Old and New Developments."

Later developments

In July 2003, the second International Conference on the Periodic Table was held in Banff (Canada).[139] A whole session was devoted to the group theoretical aspects of the periodic law, with three presentations by Novaro, Ostrovsky, and Kibler.[140] In Novaro's words, "the different approaches and conclusions of these three papers [showed] that the subject is far from exhausted."[141] Although most research on the symmetry of the periodic table was concentrated in these "schools," several other researchers have occasionally alluded to the importance of symmetry principles in understanding the periodic law.[142] The paper by Jørgensen and Katriel provides a simple but most useful graphic representation of SO(4,2) symmetry breaking, which will be a basic ingredient of our analysis in Chapter 14. Recently, L. I. Gurskiĭ and colleagues also studied the SO(4,2) group and its subsequent breaking as a tool to classify the chemical elements.[143]

13.7.2 Demkov and Ostrovsky (St. Petersburg, Russia)

All approaches reviewed so far fall short in providing a convincing explanation of Madelung's rule. The SO(4,2) group is a powerful instrument to deal with the hydrogen states; it also provides a natural splitting scheme to quantize period doubling, but, ultimately, it fails to impose the Madelung order. Other approaches, such as Novaro's, are based on artificial group extensions that do not yield a convincing explanation either. The treatment of Ostrovsky and Demkov is of a quite different nature and therefore deserves a separate section. From 1975 to 1991 Yury N. Demkov (1926–2010) was head of the Quantum Mechanics Division of the physics department in St. Petersburg. This division was founded by academician Vladimir A. Fock, who was one of the pioneers of quantum mechanics and the first to explain the hypergeometric origin of the hydrogen spectrum as discussed in Chapter 9. Valentin N. Ostrovsky (1945–2006) obtained his PhD in St. Petersburg in 1972 under the direction of Demkov. He would eventually succeed his supervisor as head of the division from 2002 until his premature death in 2006. In 1972, the year he attained

---

[139] The first International Conference on the Periodic Table celebrated the centennial of Mendeleev's discovery and was held in Turin and Rome during the third week of September 1969. The list of contributors included some of the greatest and most distinguished academicians of that time: J. P. Elliott, G. N. Flerov, M. Gell-Mann, M. Haïssinsky, L. A. Radicatti, T. Regge, E. Segré, I. Talmi, V. F. Wiesskopf, and J. A. Wheeler, to name just a few. For the conference proceedings, see M. Verde, ed. *Atti del Convegno Mendeleeviano: Periodicità e Simmetrie Nella Struttura Elementaire Della Materia*. Turin: Accademia Delle Scienze di Torino, 1971. Quite recently, in August 2012, the third International Conference on the Periodic Table was held in Cusco, Peru.

[140] See Novaro, "Group Theory of the Periodic Table"; Kibler, "A Group-Theoretical Approach to the Periodic Table: Old and New Developments"; Ostrovsky. "Group Theory Applied to the Periodic Table of the Elements."

[141] Novaro, "Symmetries of the Periodic System," p. 255.

[142] See, for example, J. Katriel and C. K. Jørgensen. "Possible Broken Supersymmetry behind the Periodic Table." *Chemical Physics Letters* 87.4 (1982), pp. 315–319; W. B. Jensen. "Classification, Symmetry and the Periodic Table." *Computers & Mathematics with Applications* 12B.1/2 (1986), pp. 487–510.

[143] L. I. Gurskiĭ, L. I. Komarov, and A. M. Solodukhin. "Group of Symmetry of the Periodic System of Chemical Elements." *International Journal of Quantum Chemistry* 72.5 (1999), pp. 499–508.

his PhD, Ostrovsky, together with Demkov, published an atomic physics approach to the Aufbau problem based on a rather peculiar family of one-electron spherically symmetric potentials:

$$V_{\text{eff}}(\mathbf{r}) = -\frac{2\nu}{r^2 R^2 \left[\left(\frac{r}{R}\right)^{-\mu} + \left(\frac{r}{R}\right)^{\mu} \right]^2}, \tag{13.45}$$

where ν, μ, and R are constant parameters.[144] This potential is different from the spherical oscillator or the Coulomb hole, and the associated Schrödinger equation cannot be integrated in closed form. Neither does a classical particle moving in this central field describe closed orbits, as Bertrand's theorem keeps reminding us (see Chapter 9). Nonetheless, this potential has some intriguing characteristics that appear only for certain specific values of the parameters. When μ is equal to one, this potential is known as the *Maxwell fish-eye potential*, with applications in geometric optics. Our interest, however, is in the field of classical and quantum mechanics. The Schrödinger equation for a particle with mass m moving in this potential thus reads

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}} \right) \Psi = E\Psi. \tag{13.46}$$

However, Demkov and Ostrovsky did not study the bound states of this equation, but instead focused on the particular solutions associated with zero energy:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}} \right) \Psi = 0. \tag{13.47}$$

Quite remarkably, for *quantized* values of the ν-parameter, this equation has analytic solutions described by Gegenbauer polynomials. The quantization condition reads

$$\nu = R^2 \mu^2 (N + \frac{1}{2\mu})(N + \frac{1}{2\mu} - 1), \tag{13.48}$$

with $N = n + ((1/\mu) - 1)l$. Obviously, the μ-value of interest here is $\mu = 1/2$, since then N reduces to $n + l$, as in the Madelung rule. For a given quantized value of N, a set of N^2 states emerges at zero energy that covers the n, l states of a given $n + l$ Madelung sum. Ostrovsky and Demkov thus leave the Schrödinger perspective and vary the potential by increasing the value of ν. Whenever ν meets the quantization condition, a Madelung manifold arises at zero energy. If ν then continues to rise, the potential deepens and these levels at zero energy move downward into the bound region. When ν passes through the next N quantum, the next set of Madelung levels appears. In this way, we obtain the Madelung spectrum, not as a function of energy, but as a function of the potential parameter. This is represented in the diagram of Figure 13.10.

Ostrovsky claims nothing less than that this result solves the Löwdin challenge, because it provides an atomic physics model that reproduces the Madelung order. The underlying rationale is that, although at larger distances the V_{eff} potential does not show the $1/r$ dependence of the Coulomb hole, its general form resembles the Thomas-Fermi potential, which provides an approximate model for multielectron atoms.

---

[144] Demkov and Ostrovsky, "$n + l$ Filling Rule in the Periodic System and Focusing Potentials."

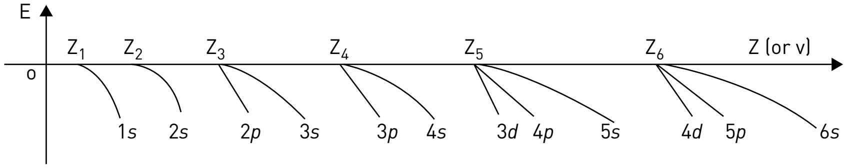


FIGURE 13.10 Schematic plot of the energy levels entering the Demkov-Ostrovsky potential well as a function of the potential strength v. Note that the s-levels ($l = 0$) are tangential to the zero-energy bound, whereas the curves for $l > 0$ cross the zero-energy line. This explains why the s-levels of a Madelung multiplet are well separated from the other levels. From V. N. Ostrovsky. "Dynamic Symmetry of Atomic Potential." *Journal of Physics B: Atomic and Molecular Physics* 14 (1981), pp. 4425–4439.

Apparently, the chemists were not really convinced by these claims. The computational chemists did not understand Löwdin's challenge in the first place, because they were able to calculate all desirable properties of any given element with great accuracy. So why bother about a general model that generates the structure of the whole periodic table? The theoretical chemists, on the other hand, were expecting an entirely different answer (i.e., in the form of a Schrödinger equation with bound states that formed a Madelung–type spectrum). Instead, they got a series of zero-energy states that were called into existence, one by one, by gradually deepening the potential well. What they also did not appreciate was that the Demkov-Ostrovsky model did retain a quantization condition. Schrödinger extracted physically sound solutions from his equation by requiring that the wave function be "finite" and single-valued, which later on would be rephrased as the requirement for square integrability of the wave function. In the Pauli's hydrogen atom, the analogue of this condition was the requirement of negative energies for the bound states. In the Demkov-Ostrovsky model, there is still a condition: only quantized values of N are considered to give rise to physically meaningful solutions.

The person who really grasped the significance of this potential, was theoretical physicist and cosmologist John Archibald Wheeler (1911–2008). In fact, his interest in the Madelung rule predates the Demkov-Ostrovsky model by three years. In 1969, at the first International Conference on the Periodic Table, Wheeler delivered a lecture titled "From Mendeleev's Atom to the Collapsing Star," in which he considered the consequences of the Madelung rule on the orbit of a classical particle inside an atom.[145] He compares and contrasts this orbit to the elliptical orbits imposed by an oscillator or Coulomb potential. The results are summarized in Table 13.9.

To explain this table we must make use of a result from Chapter 11, where it was shown that the principal quantum number in the SO(2,1) algebra could be written as a sum of a radial quantum number n_r (which corresponds to the number of nodes in the radial part of the wave function) and the angular quantum number l:

$$E(n) = E(n_r + l + 1). \tag{13.49}$$

A classical particle in a central Coulomb field describes an elliptical orbit in which the circular frequency of revolution is the same as the frequency of oscillation in the radial

---

[145] J. A. Wheeler. "From Mendeleev's Atom to the Collapsing Star." *Atti del Convegno Mendeleeviano, Accademia delle Scienze di Torino, Accademia Nazionale dei Lincei, Torino- Roma, 15–21 Settembre 1969*. Ed. M. Verde. Torino: Vincenzo Bona, 1971, pp. 189–233.

Table 13.9 The "chemical orbit" as compared to orbits in a simple harmonic oscillator and Coulomb potential.

Potential	*Orbit*	*Features of orbit preserved over entire semiclassical range of n_r and l?*	*Reaches r_{max} how many times per circuit of the orbit?*	*Passes starting direction how many times per circuit of the orbit?*
Oscillator	Centered ellipse	Yes	2	1
Coulomb	Kepler ellipse	Yes	1	1
Atom	Chemical orbit	Restricted range	1	2

From J. A. Wheeler, "From Mendeleev's Atom to the Collapsing Star." In: *Atti del Convegno Mendeleeviano, Accademia delle Scienze di Torino, Accademia Nazionale dei Lincei, Torino-Roma, 15–21 Settembre 1969*. Ed. M. Verde. Torino: Vincenzo Bona, 1971, pp. 189–233.

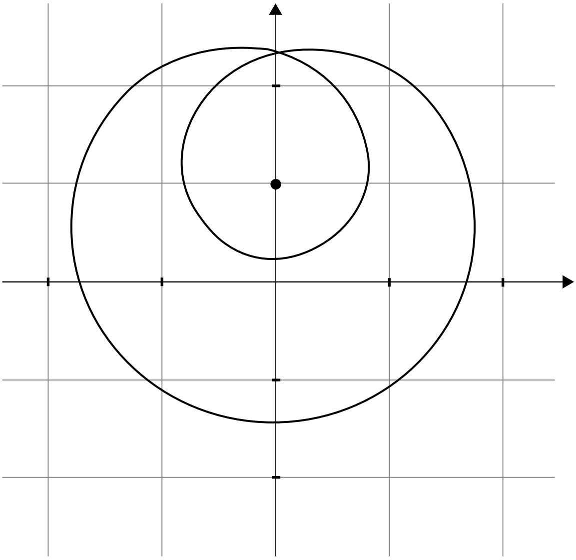


FIGURE 13.11 The double necklace, or *chemical orbit.* One complete radial oscillation from perihelion to aphelion and back involves two revolutions. From J. A. Wheeler. "From Mendeleev's Atom to the Collapsing Star." In: *Atti del Convegno Mendeleeviano, Accademia delle Scienze di Torino, Accademia Nazionale dei Lincei, Torino-Roma, 15–21 Settembre 1969*. Torino: Vincenzo Bona, 1971, pp. 189–233.

direction. Considering a single circuit around the nucleus, starting at the perihelion, it passes through the aphelion once and returns to the starting direction. In contrast, for a spherical oscillator, the energy is given by[146]

$$E = (2n_r + l + 3/2)\,\hbar\omega. \tag{13.50}$$

The corresponding orbit is a centered ellipse, hence the nucleus is located in the center between the two foci. A single circuit in this case passes through the maximal radius twice. For a particle that follows Madelung's rule, the energy is dependent on $n + l$, or

$$E(n + l) = E(n_r + 2l + 1). \tag{13.51}$$

Such an orbit in the classical approximation implies that the circular frequency of revolution is twice the frequency of excursions in the radial direction. Hence, when a particle starts (e.g., in its perihelion), it reaches the aphelion only after one full turn, and returns to the perihelion after two turns. The result is a *double necklace* trajectory, as shown in Figure 13.11. Wheeler and his assistant Robert T. Powers[147] coined the term *chemical orbit* to describe this motion. In view of Bertrand's rule, the oscillator and the Kepler problem are the only two cases with bound orbits that are closed after a single revolution. But while the rule forbids the existence of other closed, bound orbits, it cannot prevent the fact that there are some periodic orbits at peculiar energies. In 1976, Princeton University Press issued a volume of studies in mathematical physics dedicated to essays in honor of Valentine Bargmann. In that volume, Wheeler returned to the Madelung atom and took the opportunity to discuss

---

[146] For this result, see Chapter 14.

[147] R. T. Powers. "Frequencies of Radial Oscillation and Revolution as Affected by Features of a Central Potential." In: *Atti del Convegno Mendeleeviano, Accademia delle Scienze di Torino, Accademia Nazionale dei Lincei, Torino-Roma, 15–21 Settembre 1969*. Ed. M. Verde. Torino: Vincenzo Bona, 1971, pp. 235–242.

the Demkov-Ostrovsky potential.[148] With $\mu = 1/2$, the potential can be rewritten as

$$V_{\text{eff}}(\mathbf{r}) = -\frac{R(n+l+1)(n+l)}{2r(r+R)^2}. \tag{13.52}$$

Wheeler noted that this potential describes rather well the interactions that an electronic comet would feel when visiting an atomic solar system. At larger distances, the potential is determined by a polarization effect, but when the electron penetrates what Wheeler calls the *main region*, where the electron cloud becomes dense, the comet would experience a Thomas-Fermi potential that is close to the Demkov-Ostrovsky model. Deep inside the atom, close to the nucleus, the $1/r$ Coulomb potential becomes dominant, as is the case for V_{eff} with an appropriate radial scaling constant. Similarly, by slowly decreasing the value of the potential, the energy of each bound state is slowly increased until it reaches the continuum limit at the top of the atomic sea, where it is expelled into the continuum. The remarkable feature of the Demkov-Ostrovsky model is that this happens simultaneously for an entire Madelung manifold. In retrospect, by approaching the Madelung rule from a semiclassical perspective, Wheeler was able to anticipate some characteristics of the chemical orbit that were later confirmed to be present in the Demkov-Ostrovsky potential. This has contributed to the credibility of this approach. Nonetheless, it should be kept in mind that this approach focuses on the levels at zero energy and it does not describe the bound states of the atom.

Returning to Ostrovsky, in 2003 during a lecture on the physical explanation of the periodic table at a meeting of the New York Academy of Sciences, he repeated his claim to have solved the Löwdin challenge. On several occasions, he had reproached Rumer and Fet, Barut, Novaro and Berrondo, and Konopel'chenko of not having made the connection with the *inner dynamics* of many-electron systems when proposing the SO(4,2) ⊗ SU(2) group, but during the New York lecture his attack on the EPA-inspired group theoretical approaches was particularly vehement[149]:

> It seems that the abstract group-theoretical approach, as currently developed, amounts to a *translation* of empirical information on the periodicity pattern for atoms in a specialised mathematical language—but no other output is produced. Probably this approach would have explanatory power only within a community which speaks this language.

This viewpoint is reminiscent of Dirac's attitude toward the group theoretical developments that led to the eightfold way. He kept repeating that one should continue to look for Hamiltonians, no matter how difficult that might be. As Graham Farmelo recounts[150]:

> At the Lindau meeting, Dirac mounted one of his last attacks on renormalisation theory in front of an audience of some two hundred students and Nobel laureates. Looking

---

[148] J. A. Wheeler. "Semi-Classical Analysis Illuminates the Connection between Potential and Bound States and Scattering". In: *Studies in Mathematical Physics*. Eds. E. H. Lieb, B. Simon, and A. S. Wigthman. Princeton, NJ: Princeton University Press, 1976, pp. 351–422.

[149] Ostrovsky, "Physical Explanation of the Periodic Table," p. 188.

[150] G. Farmelo. *The Strangest Man: The Hidden Life of Paul Dirac, Mystic of the Atom. New York: Basic Books*, 2009, p. 405.

> as fragile as a cut-glass figurine, Dirac stood at the rostrum giving a speech almost identical to ones he had been giving for almost fifty years; he had no praise for the Standard Model or any other successes of particle physics. A microphone amplified his trembling voice, [with] each letter 's' accompanied by a whistle from his ill-fitting dentures. Current theories were "just a set of working rules" he said; physicists should go back to basics and find a Hamiltonian description of nature free from infinities. "*Some* day," he said with a gentle and weary defiance, "people will find the correct Hamiltonian."

Returning to Ostrovsky, his own negative experience with group theory might well have been at the basis of his categorical opinion on group theory. In 1981, his study of the dynamic group of the Demkov-Ostrovsky equation from an atomic physics point of view was published.[151] Interestingly, Ostrovsky's research only *revoiced* the phenomenological claim that the group theoretical aspects of the Demkov-Ostrovsky potential $O(4,2) \otimes SU(2)$ was, indeed, the symmetry group for the periodic system. Besides the usual $SU(2)$ spin group, however, Ostrovsky introduced an extra $SU(2)$ group, analogous to the isospin, to account for the period doubling. His full group thus read

$$\mathcal{G}_{\mathrm{O}} = \mathrm{O}(4,2) \otimes \mathrm{SU}(2)_{\mathrm{S}} \otimes \mathrm{SU}(2)_{\mathrm{T}}. \tag{13.53}$$

Its subgroup, $O(4) \otimes SU(2)_S \otimes SU(2)_T$, contains an $O(4)$ symmetry that yields representations of dimension n^2.[152] By enlarging this invariance group to $O(4) \otimes SU(2)_S$, the unirrep dimensions are doubled to $2n^2$. The subscript S refers to the physical origin of the $SU(2)$ group, which arises from the electron spin $m_s = \pm 1/2$. Ostrovsky referred to this "horizontal" doubling of period lengths as *spin doubling*.

The "vertical" doubling of period lengths, which is known as the actual *period doubling* in the periodic system, was given a group theoretical articulation by Ostrovsky by introducing a second $SU(2)$ group, denoted $SU(2)_T$, and was formally analogous to the isospin group. This leads to two *copies* of the $O(4,2) \otimes SU(2)_S$ unirreps, which are realized in two different Hilbert spaces.[153] The basis kets can be denoted as $|nlm_l m_s \tau\rangle$, where n, l, m_l, and m_s are defined as usual, and $\tau = \pm 1$. The value of τ then distinguishes the two Hilbert spaces from one another. To see this, Ostrovsky introduced the operators $\hat{T}_+$, $\hat{T}_-$, and $\hat{T}_3$, which were defined by the following relations[154]:

$$\hat{T}_+ \,|nlm_l m_s -\rangle \rightarrow |nlm_l m_s +\rangle\,, \quad \hat{T}_+ \,|nlm_l m_s +\rangle = 0; \tag{13.54}$$

---

[151] V. N. Ostrovsky. "Dynamic Symmetry of Atomic Potential." *Journal of Physics B: Atomic and Molecular Physics* 14 (1981), pp. 4425–4439. See also Ostrovsky, "Group Theory and Periodic System of Elements"; Ostrovsky, "What and How Physics Contributes to Understanding the Periodic Law"; Ostrovsky, "Physical Explanation of the Periodic Table"; Ostrovsky, "The Periodic Table and Quantum Physics."

[152] V. N. Ostrovsky. "Teoretiko-Gruppovye Aspekty Periodicheskoi Sistemy Elementov." In: *Group Theoretical Methods in Physics, Proceedings of International Symposium*, vol. 1. Ed. M. A. Markov. Moscow: Nauka, 1980, pp. 181–188; Ostrovsky, "Dynamic Symmetry of Atomic Potential."

[153] Although we have not adopted the dynamic group $O(4,2) \otimes SU(2)_S \otimes SU(2)_T$ in this chapter, it is interesting to note that to rationalize the Madelung $(n+l, n)$ rule, two different $\mathfrak{so}(4,2)$ copies are needed to construct the necessary Madelung operators (cf. Chapter 14).

[154] Ostrovsky, "Group Theory Applied to the Periodic Table of the Elements," p. 286.

$$\hat{T}_{-}\,|nlm_l m_s-\rangle = 0, \quad \hat{T}_{-}\,|nlm_l m_s+\rangle \to |nlm_l m_s-\rangle; \tag{13.55}$$

$$\hat{T}_3\,|nlm_l m_s\pm\rangle = \pm 1/2\,|nlm_l m_s\pm\rangle\,. \tag{13.56}$$

The $\hat{T}$ operators can then be shown to commute as

$$\left[\hat{T}_3, \hat{T}_{\pm}\right] = \pm\hat{T}_{\pm}, \tag{13.57}$$

$$\left[\hat{T}_{+}, \hat{T}_{-}\right] = 2\hat{T}_3, \tag{13.58}$$

which are the defining commutation relations for the $\mathfrak{su}(2)_{\mathrm{T}}$ algebra of the $\mathrm{SU}(2)_{\mathrm{T}}$ group. In this way, Ostrovsky was able to split the entire Hilbert space $\mathcal{H}$ into two subspaces, $\mathcal{H}_{+}$ and $\mathcal{H}_{-}$, where the $\hat{T}_3$ operator acts as a Cartan generator distinguishing states from both subspaces, and where the ladders $\hat{T}_{\pm}$ act as shift operators between $\mathcal{H}_{+}$ and $\mathcal{H}_{-}$.

Although this phenomenological construction accounts for the period doubling, it is clear in the light of Barut's discussion of the period doubling that the introduction of the $\mathrm{SU}(2)_{\mathrm{T}}$ group is excessive and unfounded physically. That is, Ostrovsky's abstract group theoretical label $\tau = \pm 1$ does not have a physical counterpart in the quantum mechanics of the periodic table and should be rejected on this basis. In contrast to the $m_s = \pm 1/2$ label, which is identified as the spin magnetic quantum number in the nonrelativistic treatment of the hydrogen atom, the label $\tau = \pm 1$ does *not* represent a *new quantum number*. It is also unclear whether a realization of the $\hat{T}_i$ operators could be obtained in terms of physical operators (expressed as analytic functions of the coordinates x_i and momenta $-i\hbar\partial/\partial x_i$).[155]

The introduction of new quantum numbers to motivate the period doubling is not an uncommon theme in the literature on the periodic table. Neubert, for example, tentatively postulated an additional quasispin property for the periodic system in terms of the *topical* quantum number $c = \pm 1/2$ to account for the *double-shell structure* of Mendeleev's chart.[156]

Extensions of the dynamic $\mathrm{O}(4,2) \otimes \mathrm{SU}(2)_{\mathrm{S}}$ group, similarly, have been proposed more than once to justify the period doubling. Following a reasoning similar to Ostrovsky's, Fet interpreted the period doubling by including the cyclic group Z_2 in the spectrum-generating group $\mathrm{O}(4,2) \otimes \mathrm{SU}(2)_{\mathrm{S}}$.[157] Although this permutation group leads to a doubling of the Aufbau scheme, it does not yield ladder operators to connect the two disjoint representations of the $\mathrm{O}(4,2) \otimes \mathrm{SU}(2)_{\mathrm{S}}$ group, as noted before by Ostrovsky.[158]

Similar objections have been raised against Konopel'chenko's postulation of the $\mathrm{SO}(2,4) + \mathrm{R}$ group, where the discrete R operator was called forth to produce the observed period doubling.[159] It should be noted that in his later publications, Fet

---

[155] Admittedly, this criticism arises from an atomic physics (APA) point of view. With the EPA, Ostrovsky's alternative construction is not as easily challenged.

[156] Neubert. "Double Shell Structure of the Periodic System of the Elements."

[157] Fet, *The System of the Elements from the Group-Theoretic Viewpoint*; A. I. Fet. "Numbers and the System of Chemical Elements." *Group Theoretical Methods in Physics, Proceedings of International Symposium*, vol. 1. Ed. M. A. Markov. Moscow: Nauka, 1980, pp. 327–336.

[158] Ostrovsky. "Group Theory Applied to the Periodic Table of the Elements," p. 279.

[159] Konopel'chenko. "Gruppa SO(2,4)+R i Tablitza Mendeleeva [The SO(2,4)+R Group and Mendeleev's Table]."

adopted the same $O(4,2) \otimes SU(2)_S \otimes SU(2)_T$ group as Ostrovsky.[160] In each of these cases, the same objections can be made as presented previously.

The final blow to Ostrovsky's findings came from A. O. Barut and Y. Kitagawara, who demonstrated that his potential well did not admit the SO(4,2) ⊗ SU(2) symmetry.[161] Although modeling the $(n+l, n)$ rule, Ostrovsky's potential did not yield an SO(4,2) symmetry. Kitagawara and Barut did not stop there, but undertook a new symmetry analysis of the Demkov-Ostrovsky potential, and dug up something more intricate and no less intriguing than SO(4,2). They discovered an algebraic structure that was quite similar to the Lie algebra SO(4), but this algebra was not closed in the strict sense. Indeed, some of the structure constants had become functions of the Hamiltonian and $\hat{L}^2$, and thus were no longer constants. Barut and Kitagawara were somewhat bewildered by this finding and wrote:

> The degeneracy algebra of this type has not been known before, and it deserves a detailed analysis.... This fact may suggest a possibility of establishing a new theory which generalizes the theory of Lie algebras. Future research will clarify these points.[162]

In a subsequent paper published in 1984, Kitagawara and Barut launched a further attempt to transform the Demkov-Ostrovsky equation into a Schrödinger–like form, without, however, changing the special algebra they had discovered. Unfortunately, their earlier promise to clarify the nature of this algebra could not be pursued.[163]

13.8 CONCLUSION

In this chapter, a group theoretical articulation was provided of the *global group structure* of the periodic system. The main results are summarized in the following four points:

1. The overall symmetry group of the bound states of the hydrogen atom has been identified with the direct product group SO(4,2) ⊗ SU(2). All possible (n, l) combinations, representing the different chemical elements, are considered to

---

[160] Fet, "The System of Elements from the Group-Theoretic Viewpoint"; Fet, "Gruppa Simmetrii Khimicheskikh Elementov" ["Symmetry Group of Chemical Elements"]; A. I. Fet and L. Romanov. "Prediction of Properties of Chemical Elements and Combinations. In: *High-Perfomance Computing and Networking*. Eds. H. Liddell, A. Colbrook, B. Hertzberger, and P. Sloot. Dordrecht: Springer, 1996, pp. 379–386; Fet, *Symmetry Group of Chemical Elements*.

[161] Y. Kitagawara and A. O. Barut. "Period Doubling in the $n+l$ Filling Rule and Dynamical Symmetry of the Demkov-Ostrovsky Atomic Model." *Journal of Physics B: Atomic and Molecular Physics* 16.18 (1983), pp. 3305–3327. See also Y. Kitagawara and A. O. Barut. "On the Dynamical Symmetry of the Periodic Table: II. Modified Demkov-Ostrovsky Atomic Model." *Journal of Physics B: Atomic and Molecular Physics* 17 (1984), pp. 4251–4259.

[162] Kitagawara and Barur. "Period Doubling in the $n+l$ Filling Rule and Dynamical Symmetry of the Demkov-Ostrovsky Atomic Model," p. 3326.

[163] In a personal communication to the authors, dated January 5, 2015, Dr. Yutaka Kitagawara recalls that, at first, the discovery of a nonlinear algebra in relation to the Madelung rule sounded odd to him. Later, he realized that the linear SO(4) algebra applicable to hydrogen was exceptional, and that one should welcome a generalized version of a Lie algebra. However, to develop such an algebra further, one should have a real physical system that can be approximated by this symmetry.

form a basis for an infinite-dimensional unirrep, denoted $h \otimes [2]$, of the SO(4,2) $\otimes$ SU(2) group.

2. A symmetry-based interpretation of the *period doubling* has been provided by Barut in terms of the reduction of the SO(4,2) group to the anti-de Sitter SO(3,2) group. The infinite-dimensional manifold of the chemical elements splits into two sets under this symmetry-breaking step: one set with $n + l$ odd and one with $n + l$ even. This leads to the required doubling of the Aufbau series, as observed in the left-step periodic system.
3. The EPA approach did not succeed in finding a group theoretical structure for the Madelung $(n + l, n)$ order. The apparent SO(4)-like pattern of the Madelung levels cannot be embedded in the SO(3,2) subgroup.
4. Demkov and Ostrovsky developed an atomic physics model that incorporates the Madelung rule, but by replacing the quantization of level energies with the quantization of coupling constants at zero energy.

Clearly, the time has come to launch a renewed EPA attack on the Madelung problem, which will be the content of our next and final chapter.

14 SO(4,2) and the rules of atomic chess

The future belongs to he who has the bishops.
–Siegbert Tarrasch[1]

14.1 FROM THE EIGHTFOLD WAY TO THE PERIODIC TABLE

In this last chapter, we finally arrive at the group theory of the periodic table. Our approach is inspired by the successful EPA that gave rise to the eightfold way. The derivation of a group structure for the periodic table will be based on three pillars:

- The period doubling;
- The Madelung rule;
- The noninvariance group SO(4,2) of the hydrogenic orbitals.

The first two are empirical observations about the periodic table; the third delineates the playing field in which the group has to be defined.

Unlike hydrogen, all heavier atoms carry more than one electron and thus incorporate an additional interelectronic repulsion energy. This effect changes the simple Coulombic central field potential responsible for the SO(4) degeneracy group of the Kepler problem (see Chapter 9). Hence, it should not be surprising that the strict hydrogenic order is no longer obeyed. In fact, the shell sequence in multielectronic

---

[1] Siegbert Tarrasch (1862–1934) was a prominent chess player and writer of the chess classic: S. Tarrasch. *The Game of Chess.* New York: Dover, 2011, p. 389. [Translated from the German: S. Tarrasch. *Das Schachspiel.* Berlin: Deutsche Buch-Gemeinschaft, 1931.]

atoms no longer only depends on the principal quantum number; it is a function of both the n and l quantum numbers. What is surprising, however, is that instead of the hydrogen sequence based on the principal quantum number, a new Madelung order seems to be imposed, which in essence remains hydrogenic, but which is based on the $n + l$ quanta. This suggests that some hitherto unknown symmetry group is taking over. The main characteristic of this symmetry group is that it traces diagonal paths through the hydrogen sequence, rather than the horizontal paths of SO(4) and the vertical paths of SO(2,1). It is precisely on this point that the existing group theoretical approaches are deficient, because they neglect the operators that correspond to diagonal moves. The iconic representation of the SO(4,2) space, which does justice to these characteristic moves, is the chessboard, as will be developed further on.

14.1.1 The atomic chessboard

Let us represent the basic nl states in a two-dimensional diagram in the form of a chessboard, with the principal quantum number n along the vertical axis and the orbital quantum number l along the horizontal axis. Since $l < n$, a triangular array is obtained (e.g., refer to Figure 14.5):

$$\begin{array}{c|ccccccc}
n\backslash l & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
7 & 7s & 7p & 7d & 7f & 7g & 7h & 7i \\
6 & 6s & 6p & 6d & 6f & 6g & 6h & \\
5 & 5s & 5p & 5d & 5f & 5g & & \\
4 & 4s & 4p & 4d & 4f & & & \\
3 & 3s & 3p & 3d & & & & \\
2 & 2s & 2p & & & & & \\
1 & 1s & & & & & &
\end{array} \tag{14.1}$$

Each entry in the table represents a manifold of $2(2l+1)$ substates, which form a tower in the third dimension. All elements belonging to a given nl configuration are thus grouped together on the same entry in the diagram. A given row (labeled by $n = 1, 2, 3, \ldots$) contains $2n^2$ elements; the columns (labeled by $l = 0, 1, 2, \ldots$) are bounded from below and contain an infinite number of elements as $n = 1 \rightarrow +\infty$. This implies that the periodic system contains (at least theoretically) an infinite number of elements; in practice, only 118 elements are known to date (with $Z = 1 \rightarrow 118$). The elements with $Z > 118$ have not yet materialized in our physical world.

The periodic table has been previously compared with a *city* with west–east *streets* in the l direction (labeled by n), and north–south *avenues* in the n direction (labeled by l). The nth street contains n *blocks*, labeled by $l = 0, 1, \ldots, n-1$. Each block, with address (n, l), contains $2(2l+1)$ *houses* with addresses given by the quartet (n, l, m_l, m_s). Chemical elements belonging to the same nl configuration are housed in the same block, and the SO(3) ⊗ SU(2) ladder operators act as *bus lines* that take us from one house to another in that same l-block. The SO(4) ⊗ SU(2) and

FIGURE 14.1 To understand the structure of the periodic system, we have to find out the rules of *atomic chess*. On the chess board in this illustration, the principal quantum number *n* increases as we walk toward the horizon. The orbital quantum number *l* increases from left to right.

SO(2,1) bus lines allow us to move along the streets and avenues, respectively, and the $\mathrm{SO}(4,2)\otimes\mathrm{SU}(2)$ *taxis* bring us from any house to any other house in Mendeleev city.[2]

However, since the relevant moves are not restricted to horizontal and vertical displacements, we will compare the diagram to a *chess board*, rather than a city map! To reconstruct the structure of the periodic table, we will have to find out the rules of this *atomic chess* (Figure 14.1).

14.1.2 Correlation diagram

The moves on a chessboard trace out the relations between the *n* and *l* quantum numbers. An overview of the significant filling rules can be presented graphically in a *correlation diagram*, as shown in Figure 14.2. This diagram goes back to an early study by Friedrich Hund (1896–1997), who drew a schematic correlation diagram between the discrete eigenvalues of the three-dimensional isotropic *harmonic oscillator* and the Schrödinger *hydrogenic levels*. The diagram was later expanded by Jörgensen and Katriel. In Figure 14.2, an n, l state is represented by a straight line, described as

$$y = n + lx. \tag{14.2}$$

---

[2] The image of streets, avenues, and houses in a city was first used by Byakov and colleagues in V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet "Group-Theoretical Classification of Chemical Elements: I. Physical Foundations." Moscow: Preprint of the Institute of Theoretical and Experimental Physics, 1976; V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. Fet "Group-Theoretical Classification of Chemical Elements: II. Description of Applied Groups." Moscow: Preprint of the Institute of Theoretical and Experimental Physics, 1976; V. M. Byakov, V. I. Kulakov, Yu. B. Rumer, and A. I. FET "Group- Theoretical Classification of Chemical Elements: III. Comparison with the Properties of Elements." Moscow: Preprint of the Institute of Theoretical and Experimental Physics 1977. The Mendeleev city metaphor has been further popularized by Maurice Kibler in a number of works. See, for example, M. R. Kibler. "On a Group-Theoretical Approach to the Periodic Table of Chemical Elements." *arXiv: quant-ph* 0408104. August 16, 2004; M. R. Kibler. "From the Mendeleev Periodic Table to Particle Physics and Back to the Periodic Table." Foundations of Chemistry 9.3 (2007), pp. 221–234.

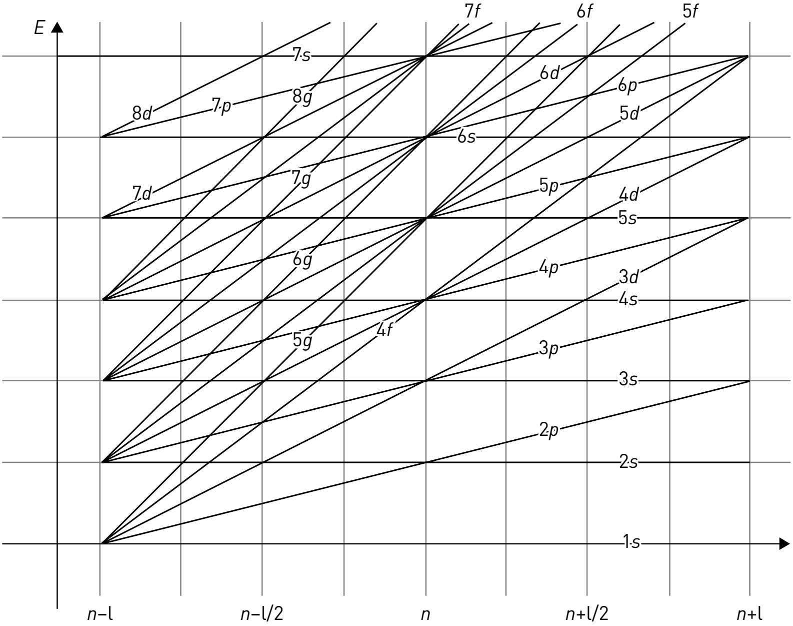


FIGURE 14.2 Qualitative correlation diagram. The energy dependence of the *nl*-orbitals on the quantum numbers *n* and *l* is depicted for five different orbital filling rules. The diagram is suggestive of a possible relation between the Madelung ($n+l$), hydrogenic (n) and harmonic ($n-l/2$) orbital orderings and degeneracies. [Adapted from J. Katriel and C. K. Jørgensen. "Possible Broken Supersymmetry behind the Periodic Table." *Chemical Physics Letters* 87.4 (1982), pp. 315–319.]

Table 14.1 Degeneracy points in the n, l correlation diagram.

x	$f_{n,l}$	*System*	*Symmetry group*
-1	$n-l$	Regge Sequence	?
$-\frac{1}{2}$	$n-\frac{l}{2}$	Isotropic oscillator	SU(3)
0	n	Hydrogen atom	SO(4)
$+\frac{1}{2}$	$n+\frac{l}{2}$	Reflected oscillator	?
$+1$	$n+l$	Madelung Sequence	?

Hence, *n* denotes the vertical scale in the origin and *l* is taken as the slope; *s* states with $l=0$ are represented by horizontal lines. This equation should not be confounded with an energy scale. It is a linearized representation of the addition of energy (or "mass" as expressed by the principal quantum number) and momentum (as expressed by the orbital quantum number) in terms of a coupling constant *x*. By changing the coupling constant from -1 to $+1$, the variation of mass and momentum can be followed along the correlation lines as the filling rule changes. The critical points correspond to integer and half-integer values of *x*, which give rise to multiple crossings between the correlation lines. At these points, different sets of *nl*-orbitals are *degenerate*, which is a hallmark of the existence of higher dynamical symmetries. In this way, the graph offers a useful chart of the possible symmetries of the central field potential. In Figure 14.2, no less than five special cases can be identified, and they are listed in Table 14.1.

Hund considered the hydrogen atom in the center of the diagram, corresponding to $x = 0$, and the harmonic oscillator system, corresponding to $n - l/2$ with $x = -1/2$. Jørgensen and Katriel later extended the diagram to the $n + l$ case, with $x = +1$, which represents the *Madelung rule*. This extension also passes through the $n + l/2$ point, which is the reflection of the oscillator. To complete the diagram, one more case has been added to the left, corresponding to the $n - l$ *anti-Madelung* (or *Regge*) rule. The Aufbau rule for monoatomic ions with charges of +2 converges back to the hydrogenic order with SO(4) symmetry in the center of the diagram. This is an important observation that confirms that an increase of the relative importance of nuclear attraction versus electronic repulsion reinforces the central Coulomb hole and diminishes the symmetry breaking due to interelectronic repulsions.

14.2 THE RULES OF ATOMIC CHESS

Each of the degeneracy spaces, identified in the correlation diagram (Figure 14.2), represents a particular move on the chessboard of the baruton, which happens to be in accordance with the traditional rules of chess. This section offers a detailed description of the different chess pieces and their associated symmetries. The allowed moves for each piece will be reviewed, along with a discussion of the corresponding ladder operators in the $\mathfrak{so}(4,2)$ Lie algebra. A summary is provided in Table 14.2.

14.2.1 The king and queen

Let us start our discussion of the chess pieces with the royals: the *king* ♔ and *queen* ♕. The king (Figure 14.3) is allowed to move one square at a time in any possible direction (horizontally, vertically, or diagonally). The queen, on the other hand, can reach out

Table 14.2 The rules of atomic chess. The symbol and allowed moves of each chess piece are summarized, along with their corresponding symmetry group and filling rule. Notice that a distinction should be made between the *black* and *white* bishops, for which $n + l$ is *even* and *odd*, respectively. The pawn is a monster; it can move upward on the chessboard only, and is thus identified with the $\hat{Q}_+$ operator. Because there is no "inverse pawn," the pawn cannot be represented by a symmetry group.

Chess piece	*Symbol*	*Moves*	*Symmetry group*	*Filling rule*
King	♔	↔ ↕ ⤢ ⤡	SO(4,2)	—
Queen	♕	↔ ↕ ⤢ ⤡	SO(4,2)	—
Rook	♖	↔ ↕	SO(4) ⊗ SO(2,1)	—
Horizontal rook	—	↔	SO(4)	n
Vertical rook	—	↕	SO(2,1)	—
Knight	♘	↩ ↪ ↪ ↩	?	$n \pm l/2$
Main knight	—	↩ ↪	SU(3)	$n - l/2$
Second knight	—	↪ ↩	?	$n + l/2$
Bishop (2x)	♗	⤢ ⤡	SO(3,2)	$n \pm l$
Regge bishop		⤢	SO(3,1)?	$n - l$
Madelung bishop		⤡	SO(4)?	$n + l$
Pawn	♙	↑	$\hat{Q}_+$	

FIGURE 14.3 The *king* ♔ can move in any possible direction by one square. It thus corresponds to the conformal SO(4,2) group.

FIGURE 14.4 The *queen* ♕ can move in all direction through any number of squares, and also corresponds to the covering SO(4,2) group.

in all directions through any number of squares (Figure 14.4). In our metaphor, the royals represent the moves allowed by the baruton itself.

14.2.2 The rook

Of greater interest is the *rook*, which will be denoted by the symbol ♖. Variously known as the *castle* or *tower*, the rook is allowed to move horizontally and vertically through any number of unoccupied squares on the board (Figure 14.5).

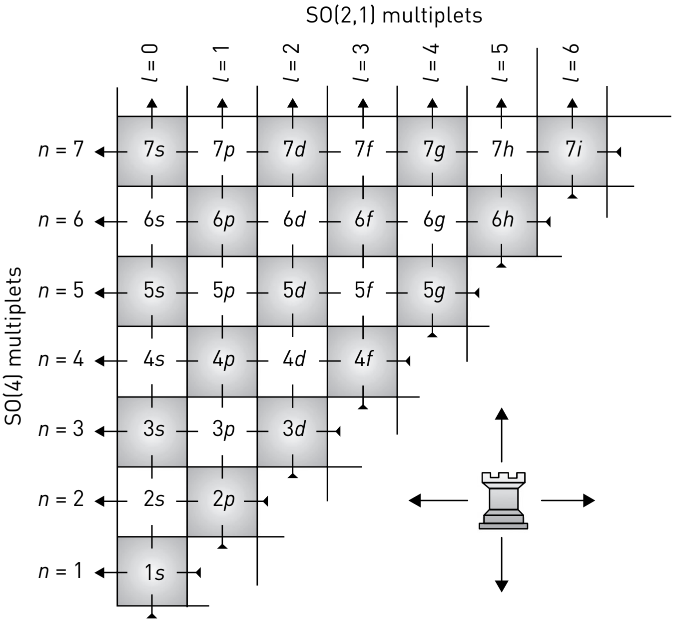


FIGURE 14.5 The *rook* ♖ can move horizontally and vertically through any number of squares. The horizontal moves yield SO(4) multiplets of constant *n*; the vertical moves give rise to SO(2,1) multiplets.

Horizontal moves of the rook on the atomic chessboard correspond to variations of l over the interval $[0, n-1]$. Orbitals with the same principal quantum number n, but different values of l, can be related in this way. The horizontal rook is therefore associated with the n rule in the center of the correlation diagram, for which the orbital energies depend uniquely on the principal quantum number n. In this sense, the rook rationalizes the accidental n^2-degeneracy of the hydrogen atom, as described by the SO(4) dynamical symmetry of the Coulomb hole (Chapter 9).

Vertical moves of the rook correspond to excitations in n, while keeping l fixed. The operators that correspond to this motion are the raising and lowering operators $\hat{Q}_\pm = \hat{Q}_1 \pm i\hat{Q}_2$ of the radial group SO(2,1).

The combined horizontal and vertical moves of the rook cover the entire chessboard and allow one to reach every possible square. This is described by the direct product group SO(4) ⊗ SO(2,1), which includes SO(4) and SO(2,1) as subgroups. It is itself a subgroup of the parent spectrum-generating group SO(4,2):

$$\mathrm{SO}(4{,}2) \;\supset\; \mathrm{SO}(4) \otimes \mathrm{SO}(2{,}1) \;\supset\; \mathrm{SO}(4), \mathrm{SO}(2{,}1). \tag{14.3}$$

The corresponding Lie algebra, $\mathfrak{so}(4) \oplus \mathfrak{so}(2,1)$ is generated by the operators $\hat{L}_i$, $\hat{A}_i$, and $\hat{Q}_i$ ($i = 1 \rightarrow 3$). An intermediate SO(4,1) group, formed by the $\hat{L}_{ab}$ operators from Chapter 12 with indices 1 to 5, is also spectrum generating. It is known as the *de Sitter group* and it contains the horizontal moves of the rook, along with the $\hat{Q}_2$ operator, which is a linear combination of the raising and lowering operators $\hat{Q}_\pm$:

$$\hat{Q}_2 = \frac{i}{2}\left(-\hat{Q}_+ + \hat{Q}_-\right). \tag{14.4}$$

This implies that $\hat{Q}_2$ turns an $|nlm\rangle$ state into a linear combination of $|(n+1)lm\rangle$ and $|(n-1)lm\rangle$ kets:

$$\begin{aligned}\hat{Q}_2\,|nlm\rangle = &-\frac{i}{2}\hbar\sqrt{(n-l)(n+l+1)}\,|(n+1)lm\rangle \\ &+\frac{i}{2}\hbar\sqrt{(n+l)(n-l-1)}\,|(n-1)lm\rangle.\end{aligned} \tag{14.5}$$

Moving away from the central point in the correlation diagram introduces a dependence on l and corresponds to the breaking of the original SO(4) symmetry. Individual energy levels originating from the hydrogen shell will now be functions of both n and l. The appropriate symmetry group can be written as the direct product group SO(3) ⊗ SO(2). The SO(3) group describes the spherical symmetry of the central field, which gives rise to the angular l quantum number, whereas the SO(2) group contains the single $\hat{Q}_3$ generator that identifies the principal quantum number n.

14.2.3 The knight

The next chess piece to be discussed is the *knight*, which will be represented by the symbol ♘ (Figure 14.6). The knight moves in an unusual manner; it can "jump" over the other pieces on the chessboard. It can move either two squares horizontally and one square vertically or two squares vertically and one square horizontally. The jumps, consequently, resemble the letter L. The intermediate $n \pm l/2$ rules are easily identified as the jumps of the knight. Although in principle the knight can reach out to eight different fields, in each of these rules only two fields on opposite sides can be addressed. The $n+l/2$ knight connects a given (n,l) state with the $(n+1,l-2)$ and $(n-1,l+2)$ states, whereas the $n-l/2$ knight jumps to the $(n+1,l+2)$ and $(n-1,l-2)$ states. As an example,

$$\begin{aligned} n+l/2 \text{ knight} \qquad & 3d \leftrightarrow 4s \\ & 4f \leftrightarrow 5p \end{aligned} \tag{14.6}$$

FIGURE 14.6 The *knight* ♘ can "jump" to one of eight possible squares by moving two squares horizontally (vertically) and one square vertically (horizontally). The knight explains the degeneracy spaces that result from the intermediate $n \pm l/2$ rules.

$$\begin{array}{ll} n-l/2 \text{ knight} & 2s \leftrightarrow 3d \\ & 3p \leftrightarrow 4f \\ & 3s \leftrightarrow 4d \leftrightarrow 5g \end{array} \tag{14.7}$$

The $n-l/2$ knight is identified as the harmonic isotropic oscillator. Since there is no group–subgroup relationship between SU(3) and SO(4), there is no direct branching rule for the correlation between the eigenstates in both systems. However, both systems share the spherical symmetry group SO(3) as largest common denominator. On this basis, we can determine the n, l quantum numbers in a given SU(3) multiplet. The angular momenta of spherical oscillator states can be worked out easily with the help of the Cartan-Weyl diagrams. The eigenstates of the Schrödinger equation for the harmonic oscillator are represented by triangulated triangles. Let N denote the number of excitations of the oscillator. The number of nodes in the triangular diagram for the N^{th} excitation is equal to $(N+1)(N+2)/2$, as we derived in Chapter 7 (see Eq. 7.73). The basis of each triangle contains $N+1$ nodes. In Figures 14.7 and 14.8 we can see the first few excitations, for $N=0,1,2,3$. As we have seen before, the $\hat{L}_z$ operator of the SO(3) subgroup of SU(3) simply corresponds to the horizontal axis in these diagrams. Hence, the m_l-values of the spherical levels can be found as projections of the nodes on the horizontal axis in this diagram. If a state with angular momentum quantum number l is present, it must give rise to a complete sequence of $2l+1$ components, with $m_l = -l, -l+1, \ldots, l-1, l$. In this way, we can immediately identify the spherical ingredients of each excitation. The SO(3) levels are labeled by their angular momentum quantum number l and a radial quantum number—say, ν—starting from 1, which enumerates the levels with the same l in increasing order, as shown in Figure 14.7. Hence, the figure reveals the following branching scheme:

$$\begin{array}{l} N=0 \rightarrow \ 1s \\ N=1 \rightarrow \ 1p \\ N=2 \rightarrow \ 2s+1d \\ N=3 \rightarrow \ 2p+1f \end{array} \tag{14.8}$$

These results can be easily generalized by considering odd and even excitations separately. The ground level ($N=0$) is a single point and thus trivially corresponds

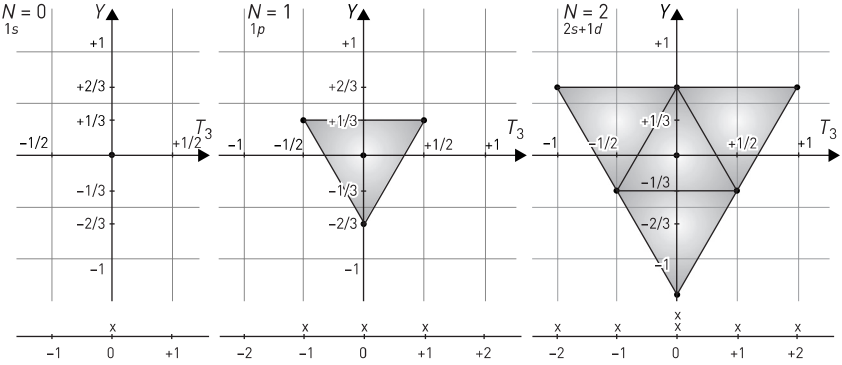


FIGURE 14.7 Cartan-Weyl diagrams for N boson excitations of the SU(3) harmonic oscillator, with $N=0,1,2$. Projections on the horizontal axis denote the m_l magnetic quantum numbers of SO(3).

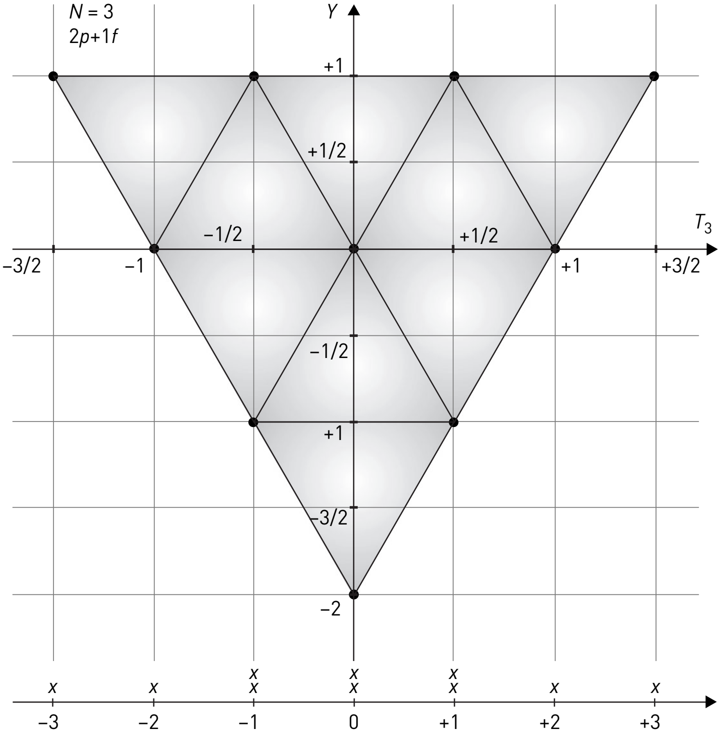


FIGURE 14.8 Cartan-Weyl diagrams for N boson excitations of the SU(3) harmonic oscillator, with $N = 3$. Projections on the horizontal axis denote the m_l magnetic quantum numbers of SO(3).

to 1*s*. The second excitation ($N = 2$) then adds to this ground level two horizontal strata, containing respectively $N + 2$ and $N + 3$ nodes. Note that the total number of nodes for this $N + 2$ excitation obeys the general dimensional formula

$$(N+2) + (N+3) + \frac{(N+1)(N+2)}{2} = \frac{[(N+2)+1]\,[(N+2)+2]}{2}. \qquad (14.9)$$

Together, the five extra points for $N = 2$ give rise to five equidistant projections on the horizontal axis, ranging from $m_l = -2$ to $m_l = +2$. This implies the existence of an extra *d*-level. We thus have a $s + d$ composition, which is labeled as $2s + 1d$, because this is the second *s*-level and the first *d*-level. Two more excitations add another two layers, now containing nine extra points, ranging from $m_l = -4$ to $m_l = +4$. The result thus corresponds to $3s + 2d + 1g$, and so on. Hence, the ground level is at the basis of a spectral series where each double excitation adds an $l = N + 2$ level, with N being even.

On the other hand, for odd values of N, the first excited state ($N = 1$) is represented by a single triangle, yielding three m_l components corresponding to the first 1*p*-level. Exactly as in the even N case, this first excited state is at the origin of a manifold, where each double excitation adds two horizontal layers, giving rise to an additional orbital with $l = N + 2$, with N being odd. Hence, 1*p* is followed by $2p + 1f$, which upon a further double excitation gives rise to $3p + 2f + 1h$ and so on.

As we saw in Chapter 7 on the oscillator, the energy of the excited states depends on N only, as it is given by

$$E(N) = \left(\frac{3}{2} + N\right)\hbar\omega. \qquad (14.10)$$

We can easily verify that there is a simple relationship between this N and the radial,[3] and angular quantum numbers n_r and l:

$$N = 2n_r + l. \tag{14.11}$$

To match these levels to the hydrogenic orbitals, we must realize that the principal hydrogenic quantum number n is defined differently. As indicated in Chapter 11, we have

$$n = n_r + l + 1. \tag{14.12}$$

Hence, the relationship between the N and n quantum numbers is given by

$$N = 2n - l - 2. \tag{14.13}$$

States belonging to an oscillator multiplet thus share the same value of $n - l/2$, which means that they are converted into each other precisely by the $n - l/2$ knight jumps.

14.2.4 The bishop

The final chess piece to pass our review is the *bishop*, denoted ♗. The bishop is allowed to move diagonally through any number of unoccupied squares on the chessboard (Figure 14.9). As a consequence, the moves of the bishop relate substates located along diagonals of the atomic chessboard. The bishop can therefore be associated with the *Madelung* $n + l$ and *anti-Madelung* $n - l$ rules. To separate these two cases, we have to define a *chiral* bishop that can move along one diagonal direction only (see §14.4).

The symmetry group that corresponds precisely to such diagonal moves is the *anti-de Sitter group*, denoted SO(3,2). This pseudo-orthogonal group of order $r = 10$ and rank $l = 2$ is a subgroup of the conformal group SO(4,2), and leaves invariant the quadratic form

$$\mathscr{F} = x_1^2 + x_2^2 + x_3^2 - x_5^2 - x_6^2, \tag{14.14}$$

which defines a five-dimensional space with signature $[+,+,+,-,-]$. The corresponding $\mathfrak{so}(3,2)$ Lie algebra is therefore based on the generalized angular momentum operators $\hat{L}_{ab}$ of Chapter 12, with indices 1, 2, 3 and 5, 6, as summarized by the matrix

$$\mathbb{L} = \begin{pmatrix} 0 & \hat{L}_{12} & \hat{L}_{13} & \hat{L}_{15} & \hat{L}_{16} \\ & 0 & \hat{L}_{23} & \hat{L}_{25} & \hat{L}_{26} \\ & & 0 & \hat{L}_{35} & \hat{L}_{36} \\ & & & 0 & \hat{L}_{56} \\ & & & & 0 \end{pmatrix}. \tag{14.15}$$

The hydrogenic realization of this group includes the $\hat{L}_i$, $\hat{B}_i$, and $\hat{\Gamma}_i$ vectors, along with the $\hat{Q}_3$ operator. Notice that the LRL vector $\hat{A}$, as well as the $\hat{Q}_1$ and $\hat{Q}_2$ components, are not included in the $\mathfrak{so}(3,2)$ algebra. Since the shift operators $\hat{Q}_\pm = \hat{Q}_1 \pm i\hat{Q}_2$ raise and lower n, whereas the $\hat{A}_i$ components alter l, horizontal and vertical moves are no longer possible under SO(3,2), and the moves are restricted to diagonals, as characteristic of the bishops.

---

[3] The radial quantum number, n_r, denotes the number of radial nodes. For $1s$, $1p$, $1d$, and so on, this number is 0.

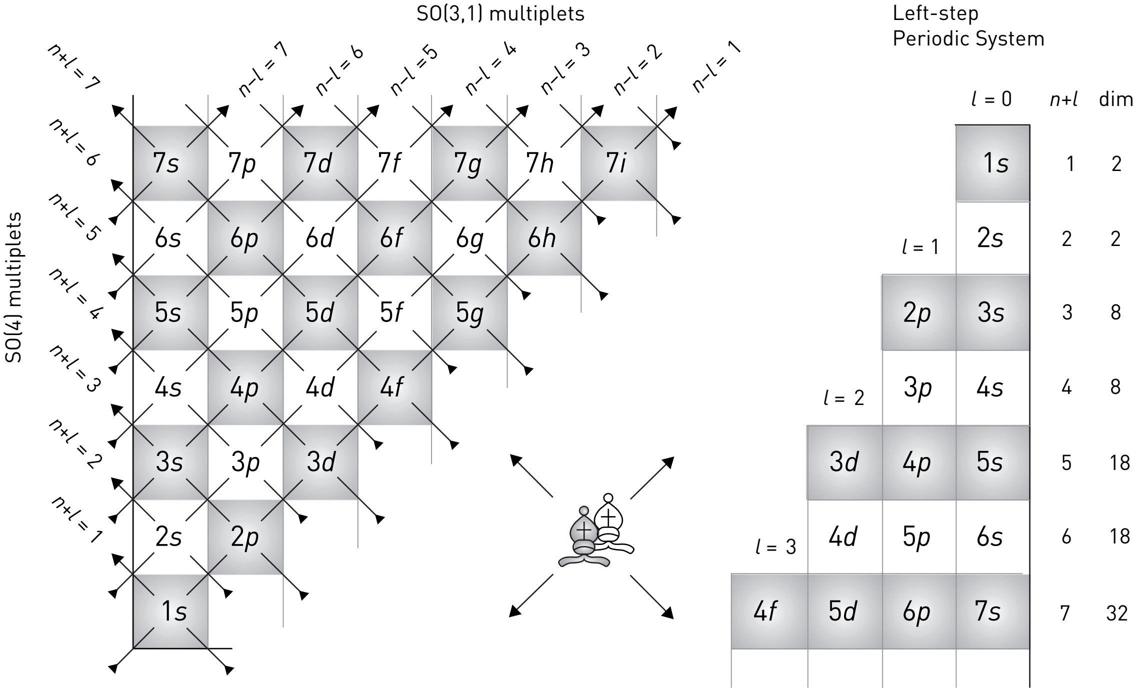


FIGURE 14.9 The *bishop* ♗ can move diagonally through any number of squares. It represents the SO(3,2) group and rationalizes the $n+l$ Madelung and $n-l$ anti-Madelung rules.

To study the moves of the bishop in greater detail, let us represent the $\hat{B}_i$ and $\hat{\Gamma}_i$ operators in ladder operator form as

$$\hat{B}_{\pm} = \hat{B}_1 \pm i\hat{B}_2; \tag{14.16}$$

$$\hat{\Gamma}_{\pm} = \hat{\Gamma}_1 \pm i\hat{\Gamma}_2. \tag{14.17}$$

The action of these operators on a ket function $|nlm\rangle$ is easily obtained from the various shift operations derived in Chapter 12. We simply need to express these operators in the known ladder operators through the following commutation rules for $\hat{B}_i$:

$$\hat{B}_3 = -\frac{1}{2}\left[\hat{Q}_+ - \hat{Q}_-, \hat{A}_3\right]; \tag{14.18}$$

$$\hat{B}_{\pm} = -\frac{1}{2}\left[\hat{Q}_+ - \hat{Q}_-, \hat{A}_{\pm}\right]; \tag{14.19}$$

and similarly for $\hat{\Gamma}_i$:

$$\hat{\Gamma}_3 = i\left[\hat{Q}_3, \hat{B}_3\right]; \tag{14.20}$$

$$\hat{\Gamma}_{\pm} = i\left[\hat{Q}_3, \hat{B}_{\pm}\right]. \tag{14.21}$$

Using the expressions from Chapter 12 leads, without further difficulty, to the explicit form of the actions of the $\hat{B}$ and $\hat{\Gamma}$ operators on a ket function $|nlm\rangle$. For the $\hat{B}$ operators,

$$\begin{aligned}
\hat{B}_3\,|nlm\rangle = {} & +\alpha^{l}_{m}u^{n}_{l}\,|(n-1)\,(l-1)\,m\rangle \\
& +\alpha^{l}_{m}v^{n}_{l}\,|(n+1)\,(l-1)\,m\rangle \\
& +\alpha^{l+1}_{m}v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,m\rangle \\
& +\alpha^{l+1}_{m}u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,m\rangle,
\end{aligned} \tag{14.22}$$

$$\begin{aligned}
\hat{B}_{\pm}\,|nlm\rangle = {} & \pm\beta^{l-1}_{\pm m}u^{n}_{l}\,|(n-1)\,(l-1)\,(m\pm 1)\rangle \\
& \pm\beta^{l-1}_{\pm m}v^{n}_{l}\,|(n+1)\,(l-1)\,(m\pm 1)\rangle \\
& \mp\gamma^{l+1}_{\pm m}v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,(m\pm 1)\rangle \\
& \mp\gamma^{l+1}_{\pm m}u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,(m\pm 1)\rangle;
\end{aligned} \tag{14.23}$$

and for the $\hat{\Gamma}$ operators,

$$\begin{aligned}
\hat{\Gamma}_3\,|nlm\rangle = {} & -i\alpha^{l}_{m}u^{n}_{l}\,|(n-1)\,(l-1)\,m\rangle \\
& +i\alpha^{l}_{m}v^{n}_{l}\,|(n+1)\,(l-1)\,m\rangle \\
& -i\alpha^{l+1}_{m}v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,m\rangle \\
& +i\alpha^{l+1}_{m}u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,m\rangle,
\end{aligned} \tag{14.24}$$

$$\begin{aligned}
\hat{\Gamma}_{\pm}\,|nlm\rangle = {} & \mp i\beta^{l-1}_{\pm m}u^{n}_{l}\,|(n-1)\,(l-1)\,(m\pm 1)\rangle \\
& \pm i\beta^{l-1}_{\pm m}v^{n}_{l}\,|(n+1)\,(l-1)\,(m\pm 1)\rangle \\
& \pm i\gamma^{l+1}_{\pm m}v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,(m\pm 1)\rangle \\
& \mp i\gamma^{l+1}_{\pm m}u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,(m\pm 1)\rangle.
\end{aligned} \tag{14.25}$$

The parameters in these expressions are defined as in Adams' *Algebraic Approach to Simple Quantum Systems*[4]:

$$\alpha^l_m = \sqrt{(l-m)(l+m)}; \tag{14.26}$$

$$\beta^l_m = \sqrt{(l-m+1)(l-m)}; \tag{14.27}$$

$$\gamma^l_m = \sqrt{(l+m+1)(l+m)} = \beta^l_{-m}; \tag{14.28}$$

$$u^n_l = \frac{1}{2}\sqrt{\frac{(n+l-1)(n+l)}{(2l-1)(2l+1)}}; \tag{14.29}$$

$$v^n_l = \frac{1}{2}\sqrt{\frac{(n-l)(n-l+1)}{(2l-1)(2l+1)}}. \tag{14.30}$$

Since the $\hat{B}_3$ and $\hat{\Gamma}_3$ operators commute with $\hat{L}_3$ (cf. Chapter 12), they do not change the m_l quantum number; so, in terms of the towers on the chessboard, we remain on the same floor. Instead, the $\hat{B}_\pm$ and $\hat{\Gamma}_\pm$ operators raise or lower the magnetic quantum number by 1 unit, in accordance with the commutation rules:

$$\left[\hat{L}_3, \hat{B}_\pm\right] = \pm\hat{B}_\pm; \tag{14.31}$$

$$\left[\hat{L}_3, \hat{\Gamma}_\pm\right] = \pm\hat{\Gamma}_\pm. \tag{14.32}$$

For a proper description of the moves of the bishop, this is an unwanted extra, which can be easily undone by preceding these operators with the inverse $\hat{L}_\pm$ ladder operators. After taking out this m_l dependence, it remains true that both $\hat{B}_\pm$ and $\hat{\Gamma}_\pm$ are sending a given state on the chessboard into a linear combination of the four adjacent states along the diagonals, as shown in Figure 14.9.

14.3 THE ORIGIN OF PERIOD DOUBLING

There is a simple but highly relevant feature to the diagonal moves. In chess, the bishop is the only *monochromatic piece*, which means it can relate fields of the same color only. For this reason, every player has two bishops to his disposition: one tied to the white fields and another one tied to the black fields. By limiting the allowed moves to diagonals, the SO(3,2) group naturally divides the baruton into two sets: one set with $n+l$ odd and one with $n+l$ even, just like the white and black fields on the chessboard, as represented in Figure 14.9. Not unlike spin, then, SO(3,2) adds an additional quantum characteristic that can take two values only: *even* and *odd*.

---

[4] B. G. Adams. *Algebraic Approach to Simple Quantum Systems.* Berlin: Springer-Verlag, 1994. See also B. G. Adams, J. Čížek, and J. Paldus. "Representation Theory of so(4, 2) for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods." *International Journal of Quantum Chemistry* 21.1 (1982), pp. 153–171; and B. G. Adams, J. Čížek, and J. Paldus. "Lie Algebraic Methods and Their Application to Simple Quantum Systems." In: *Advances in Quantum Chemistry,* vol. 19. Ed. Per-Olov Löwdin. San Diego: Academic Press, 1988, pp. 1–84, reproduced in B. G. Adams, J. Čížek, and J. Paldus. "Lie Algebraic Methods and Their Application to Simple Quantum Systems." In: *Dynamical Groups and Spectrum Generating Algebras,* vol. 1. Eds. Y. Ne'eman, A. Bohm, and A. O. Barut. Singapore: World Scientific Publishers, 1988, pp. 103–208.

The resulting splitting of the supermultiplet into two subspaces leads to a doubling of the Aufbau series. We thus arrive at the following important conclusion: *The period doubling, observed in the periodic system, is nothing more than a manifestation of a symmetry breaking from SO(4,2) to SO(3,2)*. This origin of the period doubling was noted by Barut in his lecture for the Rutherford centennial. Nonetheless, subsequently several authors have continued to postulate other mechanisms often based on the introduction of artificial spin operators.

SO(3,2) is also interesting for another reason. Although SU(3) is not a subgroup of SO(4,2), there is a group theoretical relation between SU(3) and SO(3,2). This may be seen directly in the chessboard of Figure 14.9 by leaving out either the white or the black fields. If, in the remaining halved board, we consider horizontal lines, they are seen to correspond precisely to SU(3) multiplets, as explained in the previous section. This is not a mere coincidence. Bohm has showed that the groups can be related by the operation of contraction.[5]

14.4 THE QUEST FOR THE CHIRAL BISHOP

In the correlation diagram (Figure 14.2), the diagonal moves of the bishop are found left and right of the center as the $n-l$ and $n+l$ rules respectively. Although the SO(3,2) group unites both cases, the separate cases represent a further symmetry breaking that forces the bishop to move along a single diagonal direction: either along the diagonals descending from left to right or along the diagonals ascending from left to right. On the chessboard, such moves reflect *chirality* (Figure 14.9).

14.4.1 The Madelung sequences

Diagonals descending from left to right on the chessboard link states with the same $n+l$ values, such as $5s$, $4p$, $3d$. Each diagonal forms a finite-dimensional multiplet on the board, which coincides with the periods in Janet's left-step periodic system (Figure 14.9). A study of these diagonal sequences is thus instrumental in obtaining a group theoretical interpretation of the Madelung rule. Accordingly, we will call these *Madelung sequences.*

The $n+l$ multiplets are grouped together in a single unirrep of the conformal group SO(4,2), which splits into two separate manifolds under the anti-de Sitter group SO(3,2), depending on whether $n+l$ is even or odd. The first few multiplets for both manifolds are listed in Table 14.3. Their dimensions increase as perfect squares and suggest the presence of a particular SO(4) group with multiplets that are labeled, not by the principal quantum number n, as in the case of the hydrogen atom, but by the sum of the principal and orbital quantum numbers $n+l$. However, the relationship between this SO(4) group and the SO(3,2) parent group turns out to be more complicated than we would expect.

---

[5] During the process of contraction, some Lie operators are multiplied by a numerical factor. By letting this factor tend toward zero, while keeping the resulting operators finite, some of the structure constants of the Lie algebra vanish, resulting in a contraction from SO(3,2) to SU(3). For more details, see A. Bohm, M. Loewe, P. Magnollay, M. Tarlini, R. R. Aldinger, L. C. Biedenharn, and H. van Dam. "Quantum Relativistic Oscillator: III. Contraction between the Algebras of SO(3,2) and the Three-Dimensional Harmonic Oscillator." *Physical Review* D32 (1985), pp. 2828–2834.

Table 14.3 The entire set of nl-orbitals belongs to a single unirrep of the SO(4,2) group. Under the SO(3,2) group, this manifold splits in two subsets according to whether $n+l$ is *even* (right) or *odd* (left). A further breaking of the symmetry leads to finite-dimensional multiplets with dimensions (dim) 1, 4, 9, 16, as summarized by the formula dim $= (n+l)^2$. This suggests the presence of a particular SO(4) group.

$n+l$	*Dim*	*Multiplet*	$n+l$	*Dim*	*Multiplet*
1	1	$1s$	2	1	$2s$
3	4	$3s$, $2p$	4	4	$4s$, $3p$
5	9	$5s$, $4p$, $3d$	6	9	$6s$, $5p$, $4d$
7	16	$7s$, $6p$, $5d$, $4f$	8	16	$8s$, $7p$, $6d$, $5f$

14.4.2 The Regge sequences

The diagonals ascending from left to right, the so-called *counterdiagonals*, form infinite sequences that can be viewed as excitations of *ns* ground states:

$$1s, 2p, 3d, 4f, 5g, \ldots; \tag{14.33}$$

$$2s, 3p, 4d, 5f, 6g, \ldots; \tag{14.34}$$

$$3s, 4p, 5d, 6f, 7g, \ldots. \tag{14.35}$$

We will call them *Regge sequences*, in analogy with similar hadronic sequences in elementary particle physics.[6] In 1959–1960, Tullio Regge (1931–2014) introduced the concept of Regge trajectories in scattering theory. In this theory, resonances correspond to hadronic particles, and it was soon noted that a simple linear dependence existed between the squared mass of these particles and their angular momentum:

$$m^2 = m_0^2 + aJ, \tag{14.36}$$

where a is the Regge slope. Later on, these so-called "particles" were identified as angular and radial excitations of bound di-quark or tri-quark elementary particles, corresponding to mesons and baryons, respectively. Barut and Kleinert soon realized that the sequence of linearly rising Regge states could be identified as *towers* of a SO(3,1) spectrum-generating Lorentz group.[7]

In analogy, we call the principal counterdiagonal sequence Eq. (14.33) starting from the lowest 1*s*-level in the correlation diagram a *Regge–like trajectory*. It corresponds to the line

$$n = l + 1 \tag{14.37}$$

---

[6] Consider, for instance, the weight diagrams presented by A. Bohm. "Possible Evidence for Dynamical Supersymmetry in the Hadron Spectrum." Physical Review D 33.11 (1986), pp. 3358–3367.

[7] A. O. Barut and H. Kleinert." Resonance Decays from O(3,1) Dynamics: A Regularity in the Partial Decay Widths. *Physical Review Letters*, 18, pp. 754–756. See also M. Noga and C. Cronström "Dynamical Model Leading to Linearly Rising Regge Trajectories." *Physics Letters* 29B (1969), pp. 442–444.

FIGURE 14.10 Somewhat surprisingly, Carroll never mentioned the bishop in his story of *Alice Through the Looking Glass*, perhaps out of deference to the clergy. This curious omission formed the creative source for Isaac Asimov's mystery story *The Curious Omission*. Moreover, it finds a parallel in our quest for the chiral bishop. Notice, however, that despite Carroll's failure to integrate this chess piece in the story, Tenniel did include him in one of his drawings.

and forms an infinite-dimensional supermultiplet. The parallel sequences in Eqs. (14.34) and (14.35) are called *daughter trajectories*, and they correspond to the lines

$$n = l + 2; \tag{14.38}$$

$$n = l + 3. \tag{14.39}$$

Subsequently, we will attempt to describe the spectrum-generating symmetry breaking from the SO(3,2) parent group to an SO(3,1) Regge trajectory. As we shall soon see, the relationship between this group and the SO(3,2) parent is not as simple as we would expect, and we are in for a surprise.

In summary, in order to further explain the symmetry breaking in Madelung and Regge sequences, we face the problem of *embedding* the finite diagonal sequences, such as $5s$, $4p$, $3d$, as well as the infinite diagonal sequences of hydrogenic states, corresponding to, for example, $1s$, $2p$, $3d$, $4f$, $5g$, and so on, in the SO(4,2) group. In the following sections, this quest for the *chiral bishops* will be undertaken by looking for SO(3,1) and SO(4) subgroups of the SO(4,2) group (Figure 14.10).

14.5 THE ALGEBRA OF CHIRAL BISHOPS

The potential energy function that would give rise to the Madelung or Regge sequences is unknown. So, our treatment will be purely algebraic—meaning that we start from the known expressions for the action of the SO(4,2) operators in the n, l basis and aim to derive subalgebras that correspond to the Madelung or Regge sequences. First we investigate how SO(3,2) can be reduced to an SO(3,1), which would perform the counterdiagonal moves of the Regge sequence.

14.5.1 Standard embedding of SO(3,1) in SO(3,2)

Both the $\{\hat{L}_i, \hat{B}_i\}$ and $\{\hat{L}_i, \hat{\Gamma}_i\}$ subsets form an SO(3,1) subgroup of the parent SO(3,2) and offer a natural starting point in our search for diagonal ladders. The angular momentum operators $\hat{L}_i$ and boosting operators $\hat{B}_i$ or $\hat{\Gamma}_i$ follow the standard $\mathfrak{so}(3,1)$ commutation rules:

$$\left[\hat{L}_i, \hat{L}_j\right] = i\varepsilon_{ijk}\hat{L}_k; \tag{14.40}$$

$$\left[\hat{L}_i, \hat{K}_j\right] = i\varepsilon_{ijk}\hat{K}_k; \tag{14.41}$$

$$\left[\hat{K}_i, \hat{K}_j\right] = -i\varepsilon_{ijk}\hat{L}_k, \tag{14.42}$$

where $\hat{K}_i$ stands for $\hat{B}_i$ or $\hat{\Gamma}_i$, with $i = 1,2,3$. These expressions can be reformulated in ladder operators as

$$\left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hat{L}_\pm; \tag{14.43}$$

$$\left[\hat{L}_+, \hat{L}_-\right] = 2\hat{L}_z; \tag{14.44}$$

$$\left[\hat{L}_\pm, \hat{K}_z\right] = \mp\hat{K}_\pm; \tag{14.45}$$

$$\left[\hat{K}_z, \hat{K}_\pm\right] = \mp\hat{L}_\pm; \tag{14.46}$$

$$\left[\hat{K}_+, \hat{K}_-\right] = -2\hat{L}_z. \tag{14.47}$$

Other standard embeddings of the SO(3,1) subgroup in the SO(3,2) group can be obtained by forming linear combinations of the $\hat{\mathbf{B}}$ and $\hat{\mathbf{\Gamma}}$ vectors. The diagram in Figure 14.11 shows the action of the $\hat{B}_i$ and $i\hat{\Gamma}_i$ operators. Both operators act in an identical manner, except for some sign differences, as indicated by the different arrow colors. When corresponding arrows in the $\hat{B}_i$ and $i\hat{\Gamma}_i$ diagram have the same white or black filling, it means they have the same sign; if the colors are different, they have opposite signs. The diagram thus suggests the possibility of further controlling the moves of the bishop by taking appropriate linear combinations of the $\hat{B}_i$ and $i\hat{\Gamma}_i$ components. This holds the promise of yielding new ladders that could take us up and down the Regge and Madelung trajectories.

Consider as an example the combinations $\hat{B}_\alpha \pm i\hat{\Gamma}_\alpha$, where α can be 3, +, or −. The plus combination here lowers n uniquely (Figure 14.11), whereas the minus combination can only raise n, in accordance with the commutation rule:

$$\left[\hat{Q}_3, \hat{B}_\alpha \pm i\hat{\Gamma}_\alpha\right] = \mp\left(\hat{B}_\alpha \pm i\hat{\Gamma}_\alpha\right). \tag{14.48}$$

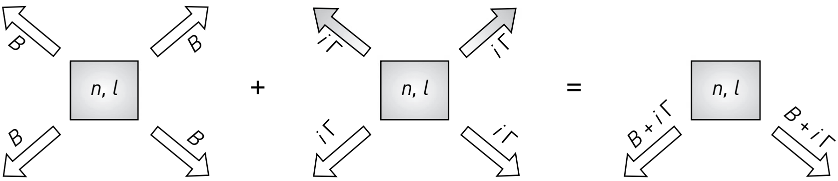


FIGURE 14.11 Additive combination of the $\hat{B}_\alpha$ and $i\hat{\Gamma}_\alpha$ operators ($\alpha = 3,+,-$) induces transitions of the central (n, l) state to $(n-1,\ l-1)$ and $(n-1,\ l+1)$ states.

The detailed expressions for the action of these operators on the standard ket are easily obtained from Eqs. (14.22) through (14.25). The plus combinations act as

$$\begin{aligned}\left(\hat{B}_3 + i\hat{\Gamma}_3\right)|nlm\rangle &= +2\alpha^l_m u^n_l \,|(n-1)(l-1)m\rangle \\ &\quad + 2\alpha^{l+1}_m v^{n-1}_{l+1}\,|(n-1)(l+1)m\rangle\,,\end{aligned} \tag{14.49}$$

$$\begin{aligned}\left(\hat{B}_- + i\hat{\Gamma}_-\right)|nlm\rangle &= -2\beta^{l-1}_{-m} u^n_l \,|(n-1)(l-1)(m-1)\rangle \\ &\quad + 2\gamma^{l+1}_{-m} v^{n-1}_{l+1}\,|(n-1)(l+1)(m-1)\rangle\,,\end{aligned} \tag{14.50}$$

$$\begin{aligned}\left(\hat{B}_+ + i\hat{\Gamma}_+\right)|nlm\rangle &= +2\beta^{l-1}_{m} u^n_l \,|(n-1)(l-1)(m+1)\rangle \\ &\quad - 2\gamma^{l+1}_{m} v^{n-1}_{l+1}\,|(n-1)(l+1)(m+1)\rangle,\end{aligned} \tag{14.51}$$

whereas the minus combinations yield

$$\begin{aligned}\left(\hat{B}_3 - i\hat{\Gamma}_3\right)|nlm\rangle &= +2\alpha^l_m v^n_l \,|(n+1)(l-1)m\rangle \\ &\quad + 2\alpha^{l+1}_m u^{n+1}_{l+1}\,|(n+1)(l+1)m\rangle\,,\end{aligned} \tag{14.52}$$

$$\begin{aligned}\left(\hat{B}_- - i\hat{\Gamma}_-\right)|nlm\rangle &= -2\beta^{l-1}_{-m} v^n_l \,|(n+1)(l-1)(m-1)\rangle \\ &\quad + 2\gamma^{l+1}_{-m} u^{n+1}_{l+1}\,|(n+1)(l+1)(m-1)\rangle\,,\end{aligned} \tag{14.53}$$

$$\begin{aligned}\left(\hat{B}_+ - i\hat{\Gamma}_+\right)|nlm\rangle &= +2\beta^{l-1}_{m} v^n_l \,|(n+1)(l-1)(m+1)\rangle \\ &\quad - 2\gamma^{l+1}_{m} u^{n+1}_{l+1}\,|(n+1)(l+1)(m+1)\rangle.\end{aligned} \tag{14.54}$$

Notice that the m dependence of the $\hat{B}_\pm \pm i\hat{\Gamma}_\pm$ operators can be taken out again by preceding their action by the inverse operator $\hat{L}_\pm$:

$$\begin{aligned}\left(\hat{B}_- + i\hat{\Gamma}_-\right)\hat{L}_+\,|nlm\rangle &= -2\omega^l_m\beta^{l-1}_{-m-1} u^n_l \,|(n-1)(l-1)m\rangle \\ &\quad + 2\omega^l_n\gamma^{l+1}_{-m-1} v^{n-1}_{l+1}\,|(n-1)(l+1)m\rangle\,,\end{aligned} \tag{14.55}$$

$$\begin{aligned}\left(\hat{B}_+ + i\hat{\Gamma}_+\right)\hat{L}_-\,|nlm\rangle &= +2\omega^l_{-m}\beta^{l-1}_{m-1} u^n_l \,|(n-1)(l-1)m\rangle \\ &\quad - 2\omega^l_{-m}\gamma^{l+1}_{m-1} v^{n-1}_{l+1}\,|(n-1)(l+1)m\rangle\,;\end{aligned} \tag{14.56}$$

$$\begin{aligned}\left(\hat{B}_- - i\hat{\Gamma}_-\right)\hat{L}_+\,|nlm\rangle &= -2\omega^l_m\beta^{l-1}_{-m-1} v^n_l \,|(n+1)(l-1)m\rangle \\ &\quad + 2\omega^l_m\gamma^{l+1}_{-m-1} u^{n+1}_{l+1}\,|(n+1)(l+1)m\rangle\,,\end{aligned} \tag{14.57}$$

$$\begin{aligned}\left(\hat{B}_+ - i\hat{\Gamma}_+\right)\hat{L}_-\,|nlm\rangle &= +2\omega^l_{-m}\beta^{l-1}_{m-1} v^n_l \,|(n+1)(l-1)m\rangle \\ &\quad - 2\omega^l_{-m}\gamma^{l+1}_{m-1} u^{n+1}_{l+1}\,|(n+1)(l+1)m\rangle\,;\end{aligned} \tag{14.58}$$

with

$$\omega^l_m = \sqrt{(l+m+1)\,(l-m)}. \tag{14.59}$$

From these results, it is also easy to check that the following *sum rule* applies:

$$2\left(\hat{B}_3 \pm i\hat{\Gamma}_3\right)\hat{L}_3 + \left(\hat{B}_+ \pm i\hat{\Gamma}_+\right)\hat{L}_- + \left(\hat{B}_- \pm i\hat{\Gamma}_-\right)\hat{L}_+ = 0. \tag{14.60}$$

This sum rule is a direct consequence of the vanishing scalar products:

$$\hat{\mathbf{B}} \cdot \hat{\mathbf{L}} = 0; \tag{14.61}$$

$$\hat{\boldsymbol{\Gamma}} \cdot \hat{\mathbf{L}} = 0. \tag{14.62}$$

This relationship also solves the apparent dilemma that we have six operators to link a given state to only four neighbors on the diagonal. In view of the sum rule, the operators, in fact, reduce to only four independent ones.

Although we have obtained one-sided combinations, they always involve a mixture of diagonal and counterdiagonal paths. It should be clear that it is impossible to form chiral ladders along either diagonal or counter diagonal directions on the chess board *unless we introduce combinations that depend explicitly on l.* The standard embedding of the SO(3,1) group in SO(3,2) does not provide an operator that yields the value of l and thus cannot yield the chiral bishops we are looking for. The same need for an additional operator that is linear in l was already noted while deriving the raising and lowering operators for l in Chapter 12.

14.5.2 The $\hat{S}$ operator

A solution in terms of a *tilted* or *distorted* SO(3,1) group will be offered in this section by introducing an additional operator $\hat{S}$, the nature of which will first be elucidated. In his book *Group Theory and the Coulomb Problem*, Englefield drew attention to an additional angular momentum operator that is not contained in the $\mathfrak{so}(4,2)$ algebra.[8] This operator, denoted $\hat{S}$, arises in the $\mathfrak{o}(3,2)$ algebra of the spherical harmonics Y_l^m and yields the first power of the angular momentum quantum number l:

$$\hat{S} Y_l^m = (l + 1/2)\, Y_l^m. \tag{14.63}$$

Notice that $\hat{S}$ acts on the angular part of the wave function $\psi = R_{n,l} Y_l^m$ only.[9] As Englefield points out, this operator is related algebraically to the angular momentum operator $\hat{L}^2$ by

$$\hat{S}^2 = \hat{L}^2 + \frac{1}{4}. \tag{14.64}$$

In fact, twice this operator yields $2l + 1$, which is the degeneracy of a given l manifold, so that $2\hat{S}$ may be considered a *counting operator*.

Although $\hat{S}$ commutes with the angular momentum operators, $\hat{\mathbf{L}}$, and radial momentum operators, $\hat{\mathbf{Q}}$, of SO(4,2), it does not commute with the LRL, $\hat{\mathbf{A}}$, nor does it commute with the remaining vector operators $\hat{\mathbf{B}}$ and $\hat{\boldsymbol{\Gamma}}$. It thus opens a way to a further *supersymmetry* beyond the original SO(4,2) symmetry. As will be shown in the following section, the $\hat{S}$ operator provides the required flexibility to obtain chiral ladders.

Let us first examine how the $\hat{S}$ operator affects the original SO(4,2) symmetry. Based on this knowledge, Regge ladder operators and Madelung ladder operators

---

[8] M. J. Englefield. *Group Theory and the Coulomb Problem*. New York: Wiley-Interscience, 1972, pp. 19–20.

[9] The other $\mathfrak{o}(3,2)$ operators are *not* symmetry operators of the hydrogen states because they shift only the spherical harmonics Y_l^m, without affecting the radial part $R_{n,l}$ simultaneously (in contrast to, for instance, the LRL vector $\hat{\mathbf{A}}$). The $\hat{S}$ operator thus forms an exception because it leaves the spherical harmonics intact, and it is also for this reason that we take the $\hat{S}$ operator onboard to see how it affects the SO(4,2) group.

can be constructed that correspond to chiral bishops. The structure of their resulting algebras will then be studied in §§14.5.4–14.5.5.

The $\hat{S}$ operator commutes with the components of $\hat{\mathbf{L}}$ and $\hat{\mathbf{Q}}$, but not with the nine other operators of the $\mathfrak{so}(4,2)$ algebra. Let us denote these non-vanishing commutators $\left[\hat{S},\hat{X}\right]$, where $\hat{X}$ represents a component of $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, or $\hat{\mathbf{\Gamma}}$, by the symbol $\hat{X}'$. Conversely, the commutator of $\hat{S}$ with one of the primed operators $\hat{X}'$ yields the original operator $\hat{X}$ again:

$$\hat{X}' = \left[\hat{S},\hat{X}\right] \quad \leftrightarrow \quad \hat{X} = \left[\hat{S},\hat{X}'\right]. \tag{14.65}$$

The primed operators $\hat{X}'$ have exactly the same action on a ket $|nlm\rangle$ as $\hat{X}$, with the only difference that moves to the left in the *nl* diagram, which lower *l*, have a different sign. The leftmost panel in Figure 14.12, for example, represents the commutator $\hat{B}' = \left[\hat{S},\hat{B}\right]$.

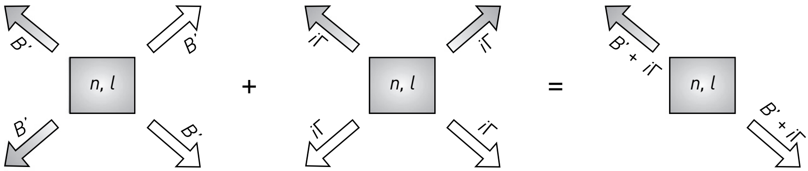
We can easily demonstrate that the commutation relations of the primed operators $\hat{\mathbf{A}}'$, $\hat{\mathbf{B}}'$, and $\hat{\mathbf{\Gamma}}'$ with operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{Q}}$ are similar to the commutators in the $\mathfrak{so}(4,2)$ algebra. In Appendix L, the commutators between the primed operators, of type $\left[\hat{X}',\hat{Y}'\right]$, are derived and we show how they can be cast in a form that makes their algebra isomorphic to the original $\mathfrak{so}(4,2)$ algebra.

The introduction of the $\hat{S}$ operator also enables us to obtain rigorous expressions for the $\hat{\Delta}_l^{\pm}$ operators introduced in Chapter 12: The original form of the $\hat{\Delta}_l^{+}$ operator was as follows:

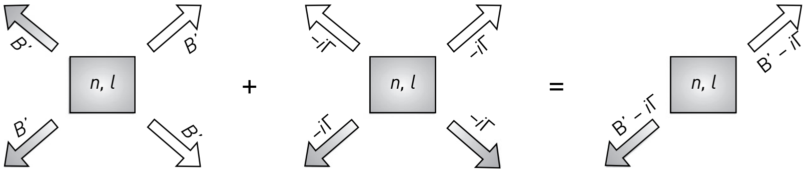
$$\hat{\Delta}_l^{+} = \frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + (l+1)\hbar\hat{A}_3. \tag{14.66}$$

We can now replace the explicit value of *l* by the $\hat{S}$ operator:

$$\hat{\Delta}_l^{+} = \frac{1}{2}\left(\hat{A}_-\hat{L}_+ - \hat{A}_+\hat{L}_-\right) + \hbar\hat{A}_3\left(\hat{S} + \frac{1}{2}\right). \tag{14.67}$$



(A) Madelung ladder operators



(B) Regge ladder operators

FIGURE 14.12 (A). Additive combination of the $\hat{B}'$ and $i\hat{\Gamma}$ operators induces diagonal transitions of the central (n,l) state to $(n+1,l-1)$ and $(n-1,l+1)$ states. (B). Additive combination of the $\hat{B}'$ and $-i\hat{\Gamma}$ operators, on the other hand, yields ladder operators for the Regge sequence.

These two expressions have exactly the same effect on the standard ket. Their action is to raise l by 1 unit, and it is given by

$$\hat{\Delta}_l^+ \, |nlm\rangle \quad = \quad (2l+1)\, \alpha_m^{l+1} c_{l+1}^n \, |n(l+1)m\rangle \, . \tag{14.68}$$

However, since we have now a full operator expression, we can define the lowering operator at once as its Hermitian conjugate, just like $\left(\hat{L}_+\right)^\dagger = \hat{L}_-$. Note that conjugation reverses the order of the products, which is indeed important because the operators $\hat{A}_3$ and $\hat{S}$ do not commute! Thus, we obtain:

$$\left(\hat{\Delta}_l^+\right)^\dagger = \frac{1}{2}\left(\hat{L}_-\hat{A}_+ - \hat{L}_+\hat{A}_-\right) + \hbar\left(\hat{S}+1/2\right)\hat{A}_3. \tag{14.69}$$

When $\hat{A}_3$ raises the l-value to $l+1$, the subsequent $\hat{S}+1/2$ operator returns the value $l+2$, which is indeed the y-coefficient preceding the $\hat{A}_3$ operator in the expression for the Δ_l^- operator in Chapter 12:

$$\hat{\Delta}_l^- = \frac{1}{2}\left(\hat{L}_-\hat{A}_+ - \hat{L}_+\hat{A}_-\right) + (l+2)\hbar\hat{A}_3. \tag{14.70}$$

So, when raising the l-value of the standard ket, both $\left(\hat{\Delta}_l^+\right)^\dagger$ and $\hat{\Delta}_l^-$ act in exactly the same way: they annihilate the ket. However, when projecting the ket to a ket with a lower l-value, there is a difference in the coefficients, because now the subsequent $\hat{S}+1/2$ operator is equal to l, compared with the fixed value of $l+2$ for the $\hat{\Delta}_l^-$ operator of Chapter 12. The results for the lowering action are

$$\begin{aligned} (\hat{\Delta}_l^+)^\dagger \, |nlm\rangle &= (2l-1)\, \alpha_m^l c_l^n \, |n(l-1)m\rangle; \\ \hat{\Delta}_l^- \, |nlm\rangle &= (2l+1)\, \alpha_m^l c_l^n \, |n(l-1)m\rangle. \end{aligned} \tag{14.71}$$

These results clearly indicate the power of the $\hat{S}$ operator, which cures the anomaly between the $\hat{\Delta}_l^\pm$ operators by providing exact Hermitian conjugation between raising and lowering operators.

14.5.3 Construction of Regge and Madelung operators

The sign alternations resulting from the commutators with $\hat{S}$ make it possible to form tilted combinations that act as ladder operators in the Regge or Madelung sequences. A convenient way to introduce these ladders is by considering the operators $\hat{Q}_3 \pm \hat{S}$, which yield the sum or difference of both the n and l quantum numbers:

$$\left(\hat{Q}_3 \pm \hat{S}\right) |nlm\rangle = (n \pm l \pm 1/2)\, |nlm\rangle. \tag{14.72}$$

We can now define the components of two new vectors, $\hat{\mathbf{R}}$ and $\hat{\mathbf{M}}$, from the commutation relations between $\hat{Q}_3 \pm \hat{S}$ and $\hat{\mathbf{B}}$. The $\hat{R}$ operators are called *Regge operators* and are given by the following set of equations:

$$\hat{R}_3 = \frac{1}{\sqrt{2}}\left[\hat{B}_3, \hat{Q}_3 + \hat{S}\right]; \tag{14.73}$$

$$\hat{R}_\pm = \frac{1}{\sqrt{2}}\left[\hat{B}_\pm, \hat{Q}_3 + \hat{S}\right]. \tag{14.74}$$

Similarly, the *Madelung operators* $\hat{M}$ are defined as

$$\hat{M}_3 = \frac{1}{\sqrt{2}}\left[\hat{B}_3, \hat{Q}_3 - \hat{S}\right]; \tag{14.75}$$

$$\hat{M}_\pm = \frac{1}{\sqrt{2}}\left[\hat{B}_\pm, \hat{Q}_3 - \hat{S}\right]. \tag{14.76}$$

The normalization factor $1/\sqrt{2}$ is introduced to obtain conformity with the structure constants in the SO(3,1) and SO(4) Lie algebras. The action of these commutators on the hydrogen states can be obtained from the preceding equations. We obtain for the Regge operators

$$\begin{aligned}\hat{R}_3\,|nlm\rangle = &-\sqrt{2}\alpha_m^{l+1}u_{l+1}^{n+1}\,|(n+1)(l+1)m\rangle \\ &+\sqrt{2}\alpha_m^{l}u_l^{n}\,|(n-1)(l-1)m\rangle\,,\end{aligned} \tag{14.77}$$

$$\begin{aligned}\hat{R}_\pm\,|nlm\rangle = &\pm\sqrt{2}\gamma_{\pm m}^{l+1}u_{l+1}^{n+1}\,|(n+1)(l+1)(m\pm 1)\rangle \\ &\pm\sqrt{2}\beta_{\pm m}^{l-1}u_l^{n}\,|(n-1)(l-1)(m\pm 1)\rangle,\end{aligned} \tag{14.78}$$

and for the Madelung operators

$$\begin{aligned}\hat{M}_3\,|nlm\rangle = &-\sqrt{2}\alpha_m^{l}v_l^{n}\,|(n+1)(l-1)m\rangle \\ &+\sqrt{2}\alpha_m^{l+1}v_{l+1}^{n-1}\,|(n-1)(l+1)m\rangle\,,\end{aligned} \tag{14.79}$$

$$\begin{aligned}\hat{M}_\pm\,|nlm\rangle = &\mp\sqrt{2}\beta_{\pm m}^{l-1}v_l^{n}\,|(n+1)(l-1)(m\pm 1)\rangle \\ &\mp\sqrt{2}\gamma_{\pm m}^{l+1}v_{l+1}^{n-1}\,|(n-1)(l+1)(m\pm 1)\rangle.\end{aligned} \tag{14.80}$$

It is clear that the Regge operators act as counterdiagonal ladders in the hydrogen spectrum, raising and lowering n and l simultaneously by 1 unit. The Madelung operators, conversely, act as diagonal ladders, and keep the sum $N = n + l$ invariant. Notice that the $\hat{\mathbf{R}}$ and $\hat{\mathbf{M}}$ vectors can also be expressed as

$$\sqrt{2}\hat{\mathbf{R}} = \left[\hat{\mathbf{B}},\ \hat{Q}_3 + \hat{S}\right] = i\hat{\Gamma} - \hat{\mathbf{B}}' = -\left(\hat{\mathbf{B}}' - i\hat{\Gamma}\right); \tag{14.81}$$

$$\sqrt{2}\hat{\mathbf{M}} = \left[\hat{\mathbf{B}},\ \hat{Q}_3 - \hat{S}\right] = i\hat{\Gamma} + \hat{\mathbf{B}}' = \hat{\mathbf{B}}' + i\hat{\Gamma}. \tag{14.82}$$

The formation of the required ladder operators thus necessitated the combination of elements from the $\mathfrak{so}(4,2)$ *and* $\mathfrak{so}'(4,2)$ algebras. This is illustrated in Figure 14.12, where the diagrams for $\hat{B}'_\alpha$ and $i\hat{\Gamma}_\alpha$ are combined according to $\hat{B}'_\alpha \pm i\hat{\Gamma}_\alpha$ (with $\alpha = 3, +, -$).

Before using these operators to construct an algebra, let us examine their Hermitian character. From Eqs. (14.77) through (14.80), it follows that both operators are anti-Hermitian. For example, for $\hat{R}_3$ we have

$$\begin{aligned}\langle (n-1)(l-1)(m)|\,\hat{R}_3\,|nlm\rangle &= \sqrt{2}\alpha_m^{l}u_l^{n} \\ \left\langle nlm|\hat{R}_3|(n-1)(l-1)(m)\right\rangle &= -\sqrt{2}\alpha_m^{l}u_l^{n}.\end{aligned} \tag{14.83}$$

Since this matrix element is real, the sign change on conjugation points out that the operator is anti-Hermitian. As another example, for the $\hat{M}_+$ operator we have:

$$\begin{aligned}\langle (n+1)(l-1)(m+1)|\,\hat{M}_+\,|nlm\rangle &= -\sqrt{2}\beta_m^{l-1}v_l^{n} \\ \left\langle nlm|\hat{M}_-|(n+1)(l-1)(m+1)\right\rangle &= +\sqrt{2}\gamma_{-m-1}^{l}v_l^{n}.\end{aligned} \tag{14.84}$$

Here, too, in view of the equality $\gamma^{l}_{-m-1} = \beta^{l-1}_{m}$, the two elements are equal, except for a sign. A similar anti-Hermitian character was noted for the prime operators, as shown in Appendix L. It can be easily reversed to Hermitian by multiplying the operators with the imaginary unit. The resulting Hermitian Regge and Madelung operators are defined as follows:

$$\begin{aligned} \hat{\mathcal{R}}_j &= i\hat{R}_j \\ \hat{\mathcal{M}}_j &= i\hat{M}_j. \end{aligned} \tag{14.85}$$

14.5.4 Regge–like algebras

In this section, the commutation relations formed by the operators $\hat{\mathbf{L}}$ and $\hat{\mathcal{R}}$ will be considered. The aim is to verify whether the set of operators $\{\hat{\mathbf{L}}, \hat{\mathcal{R}}\}$ forms an $\mathfrak{so}(3,1)$ Lorentz algebra. The commutators of type $\left[\hat{L}_i, \hat{L}_j\right]$ and $\left[\hat{L}_i, \hat{\mathcal{R}}_j\right]$ follow the standard rules for SO(3,1), since the $\hat{L}_i$ operators commute with $\hat{Q}_3 + \hat{S}$. That is,

$$\left[\hat{L}_i, \hat{L}_j\right] = i\varepsilon_{ijk}\hat{L}_k; \tag{14.86}$$

$$\begin{aligned} \left[\hat{L}_i, \hat{\mathcal{R}}_j\right] &= \frac{1}{\sqrt{2}}\left[\hat{L}_i, \left[\hat{B}_j, i\left(\hat{Q}_3 + \hat{S}\right)\right]\right] \\ &= \frac{1}{\sqrt{2}}\left[\left[\hat{L}_i, \hat{B}_j\right], i\left(\hat{Q}_3 + \hat{S}\right)\right] \\ &= i\varepsilon_{ijk}\frac{1}{\sqrt{2}}\left[\hat{B}_k, i\left(\hat{Q}_3 + \hat{S}\right)\right] \\ &= i\varepsilon_{ijk}\hat{\mathcal{R}}_k. \end{aligned} \tag{14.87}$$

The crucial point to confirm the presence of a Lorentz SO(3,1) algebra concerns the internal commutation relations of the $\hat{\mathcal{R}}_i$ components. As an example, we can work out the consecutive actions of $\hat{\mathcal{R}}_3$ and $\hat{\mathcal{R}}_{\pm}$:

$$\begin{aligned} \hat{\mathcal{R}}_3\hat{\mathcal{R}}_{\pm}\,|nlm\rangle = &+\xi_1\,|(n+2)(l+2)(m\pm 1)\rangle \\ &+\xi_2\,|nl(m\pm 1)\rangle \\ &+\xi_3\,|(n-2)(l-2)(m\pm 1)\rangle; \end{aligned} \tag{14.88}$$

$$\begin{aligned} \hat{\mathcal{R}}_{\pm}\hat{\mathcal{R}}_3\,|nlm\rangle = &+\xi_4\,|(n+2)(l+2)(m\pm 1)\rangle \\ &+\xi_5\,|nl(m\pm 1)\rangle \\ &+\xi_6\,|(n-2)(l-2)(m\pm 1)\rangle. \end{aligned} \tag{14.89}$$

The first operator raises and lowers the original (n,l) state to the immediately preceding and following states in the Regge sequence: $(n+1, l+1)$ and $(n-1, l-1)$. The second operator then takes these terms further outward to $(n+2, l+2)$ and $(n-2, l-2)$, but also projects them back to the (n,l) center. During the process, the angular momentum component is also raised or lowered by one to $(m \pm 1)$. The ξ coefficients in this expression are given by

$$\xi_1 = \pm 2\alpha^{l+2}_{m\pm 1}\gamma^{l+1}_{\pm m}\,u^{n+2}_{l+2}\,u^{n+1}_{l+1}; \tag{14.90}$$

$$\xi_2 = \pm 2\alpha^{l}_{m\pm 1}\beta^{l-1}_{\pm m}\left(u^{n}_{l}\right)^2 \mp 2\alpha^{l+1}_{m\pm 1}\gamma^{l+1}_{\pm m}\left(u^{n+1}_{l+1}\right)^2; \tag{14.91}$$

$$\xi_3 = \mp 2\alpha^{l-1}_{m\pm1}\beta^{l-1}_{\pm m}u^{n-1}_{l-1}u^n_l; \tag{14.92}$$

$$\xi_4 = \pm 2\alpha^{l+1}_{m}\gamma^{l+2}_{\pm m}u^{n+2}_{l+2}u^{n+1}_{l+1}; \tag{14.93}$$

$$\xi_5 = \mp 2\alpha^{l}_{m}\gamma^{l}_{\pm m}\left(u^n_l\right)^2 \pm 2\alpha^{l+1}_{m}\beta^{l}_{\pm m}\left(u^{n+1}_{l+1}\right)^2; \tag{14.94}$$

$$\xi_6 = \mp 2\alpha^{l}_{m}\beta^{l-2}_{\pm m}u^{n-1}_{l-1}u^n_l. \tag{14.95}$$

When substituting the coefficients in these expressions, it is gratifying to find that $\xi_1 = \xi_4$ and $\xi_3 = \xi_6$. These results are a result of the following identities:

$$\alpha^{l+2}_{m\pm1}\gamma^{l+1}_{\pm m} = \alpha^{l+1}_{m}\gamma^{l+2}_{\pm m}; \tag{14.96}$$

$$\alpha^{l-1}_{m\pm1}\beta^{l-1}_{\pm m} = \alpha^{l}_{m}\beta^{l-2}_{\pm m}. \tag{14.97}$$

This implies that when these results are combined to form the action of the commutators $\left[\hat{\mathcal{R}}_3, \hat{\mathcal{R}}_\pm\right]$, the outer terms will vanish and the commutators will only yield the $|nl(m\pm1)\rangle$ kets, which corresponds to the action of the angular momentum ladder operators $\hat{L}_\pm$. We have

$$\left[\hat{\mathcal{R}}_3, \hat{\mathcal{R}}_\pm\right]|nlm\rangle = (\xi_2 - \xi_5)\,|nl(m\pm1)\rangle. \tag{14.98}$$

This is very close to the behavior of the standard kets in the Lorentz multiplet. The factor in the equation is given by

$$\xi_2 - \xi_5 = \mp\frac{2\left(n+l+\frac{1}{2}\right)\sqrt{(l\mp m)\,(l\pm m+1)}}{2l+1}. \tag{14.99}$$

The square root in this equation is the usual prefactor of the $\hat{L}_\pm$ ladder operators, so that the commutators can be rewritten as

$$\left[\hat{\mathcal{R}}_3, \hat{\mathcal{R}}_\pm\right] = \mp\frac{\left(n+l+\frac{1}{2}\right)}{l+\frac{1}{2}}\hat{L}_\pm. \tag{14.100}$$

The coefficients in this result can be further replaced as

$$\left[\hat{\mathcal{R}}_3, \hat{\mathcal{R}}_\pm\right] = \mp\frac{\left(\hat{Q}_3+\hat{S}\right)}{\hat{S}}\hat{L}_\pm. \tag{14.101}$$

A completely analogous calculation leads to the third commutation relation:

$$\left[\hat{\mathcal{R}}_+, \hat{\mathcal{R}}_-\right] = -\frac{2\left(\hat{Q}_3+\hat{S}\right)}{\hat{S}}\hat{L}_3. \tag{14.102}$$

Clearly, these results point to a new kind of algebra. The result in the right-hand side of this equation is not the expected product of a constant structure factor and an operator. Instead, the factor preceding the $\hat{L}$ operator has become a function of $\hat{Q}_3$ and $\hat{S}$. We will analyze this unusual result in more detail later, but let us first repeat the calculation for the operators representing the Madelung rule.

14.5.5 Madelung–like algebras

The commutation relations for the operators $\hat{\mathbf{L}}$ and $\hat{\mathcal{M}}$ will be considered to verify whether this set of operators should comply with an $\mathfrak{so}(4)$ or Lorentz algebra. The commutators of type $\left[\hat{L}_i, \hat{L}_j\right]$ and $\left[\hat{L}_i, \hat{\mathcal{M}}_j\right]$ follow the standard rules for SO(4), since the $\hat{L}_i$ operators commute with $\hat{Q}_3 - \hat{S}$. That is,

$$\left[\hat{L}_i, \hat{L}_j\right] = i\varepsilon_{ijk}\hat{L}_k; \tag{14.103}$$

$$\begin{aligned} \left[\hat{L}_i, \hat{\mathcal{M}}_j\right] &= \frac{1}{\sqrt{2}}\left[\hat{L}_i, \left[\hat{B}_j, i(\hat{Q}_3 - \hat{S})\right]\right] \\ &= \frac{1}{\sqrt{2}}\left[\left[\hat{L}_i, \hat{B}_j\right], i(\hat{Q}_3 - \hat{S})\right] \\ &= i\varepsilon_{ijk}\frac{1}{\sqrt{2}}\left[\hat{B}_k, i(\hat{Q}_3 - \hat{S})\right] \\ &= i\varepsilon_{ijk}\hat{\mathcal{M}}_k. \end{aligned} \tag{14.104}$$

As before, the crucial relationships concern the internal commutation relations of the $\hat{\mathcal{M}}_i$ components. As an example, let us work out the consecutive actions of $\hat{\mathcal{M}}_3$ and $\hat{\mathcal{M}}_\pm$:

$$\begin{aligned} \hat{\mathcal{M}}_3\hat{\mathcal{M}}_\pm\,|nlm\rangle = &+\eta_1\,|(n+2)(l-2)(m\pm 1)\rangle \\ &+\eta_2\,|nl(m\pm 1)\rangle \\ &+\eta_3\,|(n-2)(l+2)(m\pm 1)\rangle; \end{aligned} \tag{14.105}$$

$$\begin{aligned} \hat{\mathcal{M}}_\pm\hat{\mathcal{M}}_3\,|nlm\rangle = &+\eta_4\,|(n+2)(l-2)(m\pm 1)\rangle \\ &+\eta_5\,|nl(m\pm 1)\rangle \\ &+\eta_6\,|(n-2)(l+2)(m\pm 1)\rangle. \end{aligned} \tag{14.106}$$

The first operator raises and lowers the original (n,l) state to the immediately preceding and following states in the Madelung sequence: $(n+1, l-1)$ and $(n-1, l+1)$. The second operator then takes these terms further outward to $(n+2, l-2)$ and $(n-2, l+2)$, but also projects them back to the (n,l) center. During the process, the angular momentum component is also raised or lowered by one to $(m\pm 1)$. The η coefficients in this expression are given by

$$\eta_1 = \mp 2\alpha^{l-1}_{m\pm 1}\beta^{l-1}_{\pm m}\nu^n_l\nu^{n+1}_{l-1}; \tag{14.107}$$

$$\eta_2 = \pm 2\alpha^{l}_{m\pm 1}\beta^{l-1}_{\pm m}\left(\nu^n_l\right)^2 \mp 2\alpha^{l+1}_{m\pm 1}\gamma^{l+1}_{\pm m}\left(\nu^{n-1}_{l+1}\right)^2; \tag{14.108}$$

$$\eta_3 = \pm 2\alpha^{l+2}_{m\pm 1}\gamma^{l+1}_{\pm m}\nu^{n-1}_{l+1}\nu^{n-2}_{l+2}; \tag{14.109}$$

$$\eta_4 = \mp 2\alpha^{l}_{m}\beta^{l-2}_{\pm m}\nu^n_l\nu^{n+1}_{l-1}; \tag{14.110}$$

$$\eta_5 = \mp 2\alpha^{l}_{m}\gamma^{l}_{\pm m}\left(\nu^n_l\right)^2 \pm 2\alpha^{l+1}_{m}\beta^{l}_{\pm m}\left(\nu^{n-1}_{l+1}\right)^2; \tag{14.111}$$

$$\eta_6 = \pm 2\alpha^{l+1}_{m}\gamma^{l+2}_{\pm m}\nu^{n-1}_{l+1}\nu^{n-2}_{l+2}. \tag{14.112}$$

When substituting the coefficients in these expressions, it is gratifying to find that $\eta_1 = \eta_4$ and $\eta_3 = \eta_6$. These results are a result of exactly the same equalities as for the Regge case (see Eqs. (14.96) and (14.97)). This implies that as these results are combined to form the action of the commutators $\left[\hat{M}_3, \hat{M}_\pm\right]$, the outer terms will vanish and the commutators will only yield the $|nl(m\pm 1)\rangle$ kets, which correspond to the action of the angular momentum ladder operators $\hat{L}_\pm$. We have

$$\left[\hat{\mathcal{M}}_3, \hat{\mathcal{M}}_+\right] |nlm\rangle = (\eta_2 - \eta_5)\, |nl(m \pm 1)\rangle. \tag{14.113}$$

This is very close to the behavior of the standard kets in the Lorentz multiplet. The factor in the equation is given by

$$\eta_2 - \eta_5 = \pm \frac{2\left(n - l - \frac{1}{2}\right)\sqrt{(l \mp m)\,(l \pm m + 1)}}{2l + 1}. \tag{14.114}$$

The square root in this equation is the usual prefactor of the $\hat{L}_\pm$ ladder operators, so that the commutators can be rewritten as

$$\left[\hat{\mathcal{M}}_3, \hat{\mathcal{M}}_\pm\right] = \pm \frac{\left(n - l - \frac{1}{2}\right)}{l + \frac{1}{2}} \hat{L}_\pm. \tag{14.115}$$

The coefficients in this result can be further replaced as

$$\left[\hat{\mathcal{M}}_3, \hat{\mathcal{M}}_\pm\right] = \pm \frac{\left(\hat{Q}_3 - \hat{S}\right)}{\hat{S}} \hat{L}_\pm. \tag{14.116}$$

A completely analogous calculation leads to the third commutation relation:

$$\left[\hat{\mathcal{M}}_+, \hat{\mathcal{M}}_-\right] = \frac{2\left(\hat{Q}_3 - \hat{S}\right)}{\hat{S}} \hat{L}_3. \tag{14.117}$$

Clearly, these results parallel the results for the Regge operators. In fact Eq. (14.116) is obtained directly from Eq. (14.101) by substituting $\hat{S}$ with $-\hat{S}$.

14.5.6 Casimir operators

The algebraic structures obtained both for the Regge and Madelung cases are close to the standard forms of SO(3,1) and SO(4) algebras. Indeed, the results can be rewritten as follows:

$$\left[\hat{L}_i,\ \hat{L}_j\right] = i\varepsilon_{ijk}\hat{L}_k, \tag{14.118}$$

$$\left[\hat{L}_i, \hat{\mathcal{R}}_j\right] = i\varepsilon_{ijk}\hat{\mathcal{R}}_k, \tag{14.119}$$

$$\left[\hat{\mathcal{R}}_i, \hat{\mathcal{R}}_j\right] = -i\varepsilon_{ijk}\zeta_R\hat{L}_k; \tag{14.120}$$

$$\left[\hat{L}_i,\ \hat{L}_j\right] = i\varepsilon_{ijk}\hat{L}_k, \tag{14.121}$$

$$\left[\hat{L}_i, \hat{\mathcal{M}}_j\right] = i\varepsilon_{ijk}\hat{\mathcal{M}}_k, \tag{14.122}$$

$$\left[\hat{\mathcal{M}}_i, \hat{\mathcal{M}}_j\right] = +i\varepsilon_{ijk}\zeta_M\hat{L}_k; \tag{14.123}$$

with

$$\zeta_R = \frac{\hat{Q}_3 + \hat{S}}{\hat{S}}$$

$$\zeta_M = \frac{\hat{Q}_3 - \hat{S}}{\hat{S}}. \tag{14.124}$$

Superficially, these resemble $\mathfrak{so}(3,1)$ and $\mathfrak{so}(4)$ Lie algebras respectively, were it not for the appearance of the ζ-factors, which depend on invariants of the subalgebras. The algebraic structure we have obtained is said to be *nonlinear* in the sense that the structure factors are *functions* rather than *constants.*[10] This nonlinearity also implies an alteration of the expressions for the Casimir operators. Let us first examine this for the Regge operators.

Casimir operators for the Regge algebra

The ζ_R structure constant for the Regge algebra can be rewritten as follows:

$$\zeta_R = \frac{\hat{Q}_3 + \hat{S}}{\hat{S}} = a_{-1}\left(\frac{1}{S}\right) + a_0, \tag{14.125}$$

where $a_{-1} = \hat{Q}_3 - \hat{S}$ and $a_0 = 2$. The parameter a_{-1} is a constant for a given Regge–like sequence. As an example, in the *fundamental sequence* $1s, 2p, 3d, 4f, 5g$, and so on, this factor equals $1/2$. In the *daughter sequence* $2s, 3p, 4d, 5f, 6g$, and so on, it equals $3/2$, and so on. The nonlinear $\mathfrak{so}(3,1)$ algebra which we have obtained is *associative* and obeys the *Jacobi identity*. Following Racah's theorem and the fact that we are dealing with an algebra of order $r = 6$ and rank $l = 2$, we expect to find two Casimir operators $\hat{C}_\mu$ that commute with all the generators $\hat{L}_i$ and $\hat{\mathcal{R}}_i$. One of these is straightforward:

$$\hat{C}_2 = \hat{\mathbf{L}} \cdot \hat{\mathcal{R}} = 0. \tag{14.126}$$

The other Casimir invariant requires the use of the same scaling factor ζ_R as appeared in the commutation relationships. First, we calculate the length of the $\hat{\mathcal{R}}$ vector:

$$\hat{\mathcal{R}}^2 = \hat{\mathcal{R}}_3^2 + \frac{1}{2}\left(\hat{\mathcal{R}}_+\hat{\mathcal{R}}_- + \hat{\mathcal{R}}_-\hat{\mathcal{R}}_+\right). \tag{14.127}$$

The products of $\hat{\mathcal{R}}$ operators turn the starting $|nlm\rangle$ ket into doubly raised and lowered $(n \pm 2, l \pm 2)$ combinations, but also project it back onto itself. It can be easily shown that the resulting coefficients of the doubly excited and annihilated states vanish, as we observed for the commutators in the previous section. The diagonal coefficients for the (n,l) state are as follows:

$$\begin{aligned}
\langle nlm|\,\hat{\mathcal{R}}_3^2\,|nlm\rangle &= 2\left(\alpha^l_m\right)^2\left(u^n_l\right)^2 + 2\left(\alpha^{l+1}_m\right)^2\left(u^{n+1}_{l+1}\right)^2\\
\langle nlm|\,\hat{\mathcal{R}}_+\hat{\mathcal{R}}_-\,|nlm\rangle &= 2\beta^{l-1}_{-m}\beta^l_{-m+1}\left(u^n_l\right)^2 + 2\beta^{l+1}_m\beta^l_{m-1}\left(u^{n+1}_{l+1}\right)^2\\
\langle nlm|\,\hat{\mathcal{R}}_-\hat{\mathcal{R}}_+\,|nlm\rangle &= 2\beta^{l-1}_m\beta^l_{m+1}\left(u^n_l\right)^2 + 2\beta^{l+1}_{-m}\beta^l_{-m-1}\left(u^{n+1}_{l+1}\right)^2.
\end{aligned} \tag{14.128}$$

---

[10] A *nonlinear deformation* of a Lie algebra has generators $\hat{T}_i$ for which the commutator $\left[\hat{T}_i, \hat{T}_j\right]$ is not expressed as a linear combination of the generators $\hat{T}_k$, but as a sum over terms $f_{ijk}\left(\hat{T}_l\right)\hat{T}_k$, where f_{ijk} are polynomials in the generators $\hat{T}_l$:

$$\left[\hat{T}_i, \hat{T}_j\right] = \sum_k f_{ijk}\left(\hat{T}_l\right)\hat{T}_k.$$

See, for example, J. Fuchs and C. Schweigert. "Open and Non-linear Algebras." In: *Symmetries, Lie Algebras and Representations: A Graduate Course for Physicists.* Cambridge: Cambridge University Press, 1997, pp. 60–61.

Combining these results in accordance with Eq. (14.127) yields

$$\langle nlm|\hat{\mathcal{R}}^2|nlm\rangle = \frac{l(n+l)(n+l-1)}{2(2l+1)} + \frac{(l+1)(n+l+1)(n+l+2)}{2(2l+1)} \tag{14.129}$$

In operator form, this can be rewritten as

$$\hat{\mathcal{R}}^2 = \frac{3}{8} + \frac{1}{2}\frac{\hat{Q}_3+\hat{S}}{\hat{S}} + \frac{1}{2}\left(\hat{Q}_3+\hat{S}\right)^2. \tag{14.130}$$

The nonlinearity of the algebra requires us to look for a modified Casimir invariant $\hat{C}_1$ of the form $f\left(\hat{L}^2\right) - \hat{\mathcal{R}}^2$, where

$$f\left(\hat{L}^2\right) = \left(\frac{\hat{Q}_3+\hat{S}}{\hat{S}} - 1\right)\left(1+2\hat{L}^2\right). \tag{14.131}$$

The first Casimir operator then becomes

$$\hat{C}_1 = f\left(\hat{L}^2\right) - \hat{\mathcal{R}}^2 = -\frac{7}{8} - \frac{1}{2}\left(\hat{Q}_3-\hat{S}\right)^2. \tag{14.132}$$

As we saw earlier, the Regge sequences are characterized by a constant value for $\hat{Q}_3 - \hat{S}$, and thus are eigenfunctions of the $\hat{C}_1$ operator. It is easily verified that the eigenvalue of this operator is a negative integer. For the principal series, starting at $1s$, $\hat{C}_1$ equals -1; for the daughter series, starting at $2s$, $\hat{C}_2$ equals -2. In general, for the nth series, starting at ns, $\hat{C}_1$ is given by:

$$\hat{C}_1 = -\left(\frac{n^2-n}{2}+1\right). \tag{14.133}$$

Casimir operators for the Madelung algebra

An analogous treatment can also be pursued for the Madelung operators $\hat{\mathcal{M}}_i$. The ζ_M factor can be rewritten as follows:

$$\zeta_M = \frac{\hat{Q}_3-\hat{S}}{\hat{S}} = b_{-1}\left(\frac{1}{S}\right) + b_0, \tag{14.134}$$

where $b_{-1} = \hat{Q}_3 + \hat{S}$ and $b_0 = -2$. The parameter b_{-1} is a constant for a given Madelung sequence. As an example, in the sequence $4f, 5d, 6p, 7s$, this factor equals $15/2$. In the sequence $4d, 5p, 6s$, it equals $13/2$, and so on. The nonlinear $\mathfrak{so}(4)$ algebra we obtained is *associative* and obeys the *Jacobi identity*. Following Racah's theorem and the fact that we are dealing with an algebra of order $r = 6$ and rank $l = 2$, we expect to find two Casimir operators $\hat{C}_\mu$ that commute with all the generators $\hat{L}_i$ and $\hat{\mathcal{M}}_i$. One of these is straightforward:

$$\hat{C}_2 = \hat{\mathbf{L}} \cdot \hat{\mathcal{M}} = 0. \tag{14.135}$$

The other Casimir invariant requires the use of a similar scaling factor ζ as appeared in the commutation relationships. First, we calculate the length of the $\hat{\mathcal{M}}$ vector:

$$\hat{\mathcal{M}}^2 = \hat{\mathcal{M}}_3^2 + \frac{1}{2}\left(\hat{\mathcal{M}}_+\hat{\mathcal{M}}_- + \hat{\mathcal{M}}_-\hat{\mathcal{M}}_+\right). \tag{14.136}$$

As for the Regge case, this operator turns the starting $|nlm\rangle$ ket onto itself. The coefficients for the (n,l) state are as follows:

$$\begin{aligned}
\langle nlm|\,\hat{M}_3^2\,|nlm\rangle &= 2\left(\alpha^{l}_{m}\right)^2\left(v^{n}_{l}\right)^2+2\left(\alpha^{l+1}_{m}\right)^2\left(v^{n-1}_{l+1}\right)^2\\
\langle nlm|\,\hat{M}_+\hat{M}_-\,|nlm\rangle &= 2\beta^{l-1}_{-m}\gamma^{l}_{m-1}\left(v^{n}_{l}\right)^2+2\beta^{l}_{m-1}\gamma^{l+1}_{-m}\left(v^{n-1}_{l+1}\right)^2\\
\langle nlm|\,\hat{M}_-\hat{M}_+\,|nlm\rangle &= 2\beta^{l-1}_{m}\gamma^{l}_{-m-1}\left(v^{n}_{l}\right)^2+2\beta^{l}_{-m-1}\gamma^{l+1}_{m}\left(v^{n-1}_{l+1}\right)^2.
\end{aligned} \tag{14.137}$$

Combining these results in accordance with Eq. (14.127) yields

$$\begin{aligned}
\langle nlm|\,\hat{\mathcal{M}}^2\,|nlm\rangle &= \frac{l(n-l)(n-l+1)}{2(2l+1)}+\frac{(l+1)(n-l-1)(n-l-2)}{2(2l+1)}\\
&= \frac{1}{2}(n-l-1/2)^2-\frac{n-l-1/2}{2l+1}+\frac{3}{8}.
\end{aligned} \tag{14.138}$$

In operator form, this can be rewritten as

$$\hat{\mathcal{M}}^2=\frac{3}{8}-\frac{1}{2}\frac{\hat{Q}_3-\hat{S}}{\hat{S}}+\frac{1}{2}\left(\hat{Q}_3-\hat{S}\right)^2, \tag{14.139}$$

where we have taken into account the rescaling in Eqs. (14.121),(14.122), and (14.123). The nonlinearity of the algebra requires us to look for a modified Casimir invariant $\hat{C}_1$ of the form $f\left(\hat{L}^2\right)+\hat{\mathcal{M}}^2$, where

$$f\left(\hat{L}^2\right)=\left(\frac{\hat{Q}_3+\hat{S}}{\hat{S}}-1\right)\left(1+2\hat{L}^2\right). \tag{14.140}$$

The first Casimir operator then becomes

$$\hat{C}_1=f\left(\hat{L}^2\right)+\hat{\mathcal{M}}^2=\frac{7}{8}+\frac{1}{2}\left(\hat{Q}_3+\hat{S}\right)^2. \tag{14.141}$$

As we have seen, the Madelung sequences are characterized by a constant value for $\hat{Q}_3+\hat{S}$, and thus are eigenfunctions of the $\hat{C}_1$ operator. It can be easily proved that as a result of the applied scaling, the eigenvalue of this operator is a positive integer. For the principal series, starting at 1*s*, $\hat{C}_1=2$; for the daughter series, starting at 2*s*, $\hat{C}_2=4$. In general, for the *n*th series, starting at *ns*, $\hat{C}_1$ is given by

$$\hat{C}_1=\left(\frac{n^2+n}{2}+1\right). \tag{14.142}$$

14.6 NONLINEAR LIE ALGEBRAS

The commutation of the $\hat{Q}_3$ operator with the diagonal $\hat{B}_i$ operators introduces into the chessboard of hydrogenic states an upper/lower reflection whereas the additional $\hat{S}$ operator is responsible for a left/right reflection (Figures 14.11, 14.12). The combination of both is needed to generate sequences that run along the diagonals. The resulting Regge and Madelung algebras are peculiar in the sense that they are *nonlinear*, while maintaining the angular momentum algebra as an $\mathfrak{so}(3)$ subalgebra.

Various similar extensions of the traditional Lie algebraic framework have started to play an increasingly important role in many branches of quantum physics. Quantum algebras (such as *q*-deformed Lie [super]algebras) and other quantum deformations with nonlinear structure functions are increasingly applied to many-body quantum physics and hold the promise of shedding new light on their internal dynamics.[11]

The algebras of the chiral bishops bear a close resemblance to the nonlinear extensions of the $\mathfrak{so}(3,1)$ and $\mathfrak{so}(4)$ algebras that have been discussed by Quesne.[12] Quesne considered deformations in which the structure constants are power series in the group generators, fulfilling the associativity and Jacobi requirements, and that contract to a linear algebra for some limiting values of the parameters. Compared with the power series law considered by Quesne, the current structure functions present an inverse power dependence on the angular momentum through the ζ_R and ζ_M factors in Eq. (14.124). So far, such a case has not yet been discussed in the specialized literature.

In their analysis of the Demkov-Ostrovsky potential, as described in Chapter 13, Barut and Kitagawara stumbled onto a "new algebra" that resembled the $\mathfrak{so}(4)$ algebra, but was not closed in the strict sense. This was in 1983, at a time when the importance of non-linear Lie algebra's was not yet recognized, and the expressions were quite complicated because the Casimir operators involved the fish-eye potential itself. In contrast, the current treatment is focused uniquely on the EPA approach, and yet arrives at the converging result that the Madelung order requires a distortion of the hydrogenic SO(4) symmetry by a non-linearity. The expressions are more simple than the ones of Barut and Kitagawara because they contain only the SO(4,2) operators and their commutators with the $\hat{S}$ operator.

14.7 CONCLUSION

In this chapter, a group theoretical articulation was provided of the *global group structure* of the periodic system. The main results can be summarized in the following three points:

1. The overall symmetry group of the bound states of the hydrogen atom is given by the direct product group SO(4,2) $\otimes$ SU(2), where SU(2) accounts for the spin part. All possible (n,l) combinations, representing the different chemical elements, are considered to form a basis for an infinite-dimensional unirrep, denoted $h \otimes [2]$, of the SO(4,2) $\otimes$ SU(2) group.
2. A symmetry-based interpretation of the *period doubling* is provided in terms of the reduction of the SO(4,2) group to the anti-de Sitter SO(3,2) group. The infinite-dimensional manifold of the chemical elements splits into two sets under this symmetry-breaking step: one set with $n+l$ being odd and one with $n+l$ being even. This leads to the required doubling of the Aufbau series, as observed in the left-step periodic system.

---

[11] See, for instance, M. R. Kibler and T. Negadi. "On the q-Analogue of the Hydrogen Atom." *Journal of Physics A: Mathematical and General* 24.22 (1991), pp. 5283–5289; M. R. Kibler and T. Negadi. "A q-Deformed Aufbau Prinzip." *Journal of Physics A: Mathematical and General 25.4 (1992)*, pp. L157–L160; T. Negadi and M. R. Kibler. "The Periodic Table in Flatland." *International Journal of Quantum Chemistry* 57.1 (1996), pp. 53–61; T. Negadi. "On the Planar Periodic Table." International Journal of Quantum Chemistry 78.4 (2000). pp. 206–211.

[12] C. Quesne. "On Some Nonlinear Extensions of the Angular Momentum Algebra." *Journal of Physics* A: Mathematical and General 28.10 (1995), pp. 2847–2860.

3. The Madelung $(n+l, n)$ rule was described in a group theoretical manner by invoking an $\hat{S}$ operator that yields a new $\mathfrak{so}'(4,2)$ algebra under commutation with the original $\mathfrak{so}(4,2)$. This enabled the construction of Madelung operators $\hat{M}_i$ by combining generators from both algebras. The $\hat{M}_i$ commute with the angular momentum vectors $\hat{L}_i$ to form a *nonlinear* $\mathfrak{so}(4)$ algebra. The nonlinearity of this algebra reflects the screening of the Coulomb potential in many-electron systems.

Although repulsion modifies the hydrogen potential, and thus breaks the SO(4) symmetry, the truly fascinating aspect of the sequences in the periodic table, is that, after all, the periodicity continues to respect the magic number series of SO(4), but now traces diagonal paths instead of horizontal ones. This altered but still hydrogenic structure finds its group theoretical expression in the replacement of the SO(4) algebra by a nonlinear congener.

Epilogue

Mathematics is about existence. A regular polyhedron is defined as a closed body for which all vertices, all edges, and all faces are the same. The doctrine of the Platonic solids claims that in three dimensions there exist five regular polyhedra—and only five. The fact that no other solutions can exist did not put the Greek mathematicians to rest, though. By sacrificing the requirement that all faces be the same, a new family of regular polyhedra was called into existence: the Archimedean solids, of which there are exactly thirteen. They include the truncated icosahedral structure, which is realized in Buckminsterfullerene, and the cuboctahedron, which is the root diagram of SO(4,2). As this history shows, the judicious removal of a specific constraint, without doing harm to the core concept, can lead to new discoveries.

Similar tricks were played in number theory. In antiquity, the Pythagoreans as well as Indian mathematicians were aware of the existence of irrational numbers that could not be expressed as the ratio of integers. The resulting set of the real numbers marks an end point of number theory. It almost put the mathematics of number systems to rest. But, after a long dormant period, the number theory was "reborn" during the Renaissance with the discovery of complex numbers. The change of paradigm that triggered this discovery was to give up the implicit requirement that numbers be one-dimensional. Complex numbers have a real and an imaginary component; hence, they live on a two-dimensional plane but nevertheless form an algebra (i.e., they can be added, multiplied, and divided, and they can be characterized by a "length," or norm, such that the norm of a product of two complex numbers is the product of their norms). When the dimensional constraint had fallen, the hunt was on for number systems in more dimensions. In 1843, Hamilton, on his famous walk, came to meet the quaternions: a number system with four components. They, too, form a so-called *division algebra*, but only at the expense of giving up commutativity of the product of quaternions. One more number system remained to be discovered: the so-called *octonions*, or *Cayley numbers*, which have eight components. In this case, the requirement of associativity had to surrender. And then, in 1898, Hurwitz published his famous theorem that, with this final surrender, the list of solutions of these normed algebras was complete and no other such algebras could exist. A century-long program had reached its completion.[13]

Also, group theory has known its own heroic history. Finite groups can be broken up in so-called *simple groups* much in the same way as numbers can be decomposed into products of prime numbers. The pioneering work of Galois and Lie had laid the foundation for the classification of simple groups in a periodic system of a kind. Throughtout the twentieth century, this periodic system continued to grow. Would this growth continue forever? Or

---

[13] F. Gürsey. "Quaternionic and Octonionic Structures in Physics. Episodes in the Relation between Physics and Mathematics." In: *Symmetries in Physics (1600–1980), First International Meeting on the History of Scientific Ideas (Sant Feliu de Guixols, Catalunya, September 20–26, 1983)*. Eds. M. G. Doncel, A. Hermann, L. Michel, A. Pais. Barcelona: University Press, 1987, pp. 557–592.

would it reach completion? The final verdict came during the 1980s. The hunt for new classes of finite groups stopped; the classification scheme was complete. This conclusion is the result of a worldwide effort of mathematicians and ran over tens of thousands of pages spread throughout the mathematical literature, with echoes extending until 2008. It terminates a long period of intense work on group structures, but at the same time opens the gate to new challenges.[14]

In this book, which we now leave behind, a similar story of opening and closing was read. As mentioned in Part II, Joseph Bertrand showed, in 1873, that in classical mechanics there exist two, and only two, central force potentials that give rise to closed bound orbits: the inverse square law of the gravitation and electrostatics, and Hooke's law of the radial oscillator. Much of the success of quantum mechanics is a result of the omnipresence of these potentials in physics and chemistry: the radial oscillator for the quark model and the inverse square law for the hydrogen atom. Both cases are marvelously described by Lie algebras, as we saw in Chapters 7 and 9. They form, in a sense, the apogee of Lie algebra, but at the same time also show its limits. If it comes to the periodic table, the hydrogenic order breaks down. The potential becomes more involved as a result of interelectronic repulsion, and it is natural to claim a breakdown of SO(4) symmetry, leaving behind a shattered Lie symmetry. But, this is not the full story. As we saw in the final chapters, a new hydrogenic order is installed, which is described as a tilting across the (n, l) chessboard, corresponding to the diagonal moves of the bishops. We derived the corresponding operators and observed that they form a nonlinear Lie algebra. Giving up the requirement of linearity should not be deplored as a loss, but welcomed as a new opportunity.

---

[14] M. Du Sautoy. *Finding Moonshine: A Mathematician's Journey through Symmetry*. London: Fourth Estate, Harper Collins Publishers, 2008.

Appendix A
Vector algebra

A.1 CONCEPTUAL DEFINITION

Some physical quantities, such as mass, volume, density, length, temperature, energy, and charge, are determined entirely by their magnitude and possibly a sign. They are called *scalars*. Physical quantities that have both magnitude *and* direction are termed *vectors*. Examples include velocity, acceleration, force, linear momentum, and angular momentum.

A.1.1 Notation

Vectors are often represented by a directed line segment (or an arrow) that connects an initial point A with a final point B. The vector's magnitude and direction are then specified by the length and orientation of the line segment. Line segments are often denoted as $\overrightarrow{AB}$, with A the tail and B the head of the arrow. In this book, however, all vectors are denoted by lowercase boldface symbols, such as $\mathbf{a}$. The length or magnitude of a vector $\mathbf{a}$ is denoted by the symbol $|\mathbf{a}|$.

A.2 REPRESENTATION

Vectors are commonly represented in a Cartesian coordinate system. Consider a vector $\mathbf{a}$ in three-dimensional Euclidean space $\mathbb{R}^3$ with a tail that coincides with the origin $O = (0,0,0)$ of a Cartesian $\Sigma\left(x,y,z\right)$ system, and with an end point that is given by the set of coordinates $\left(a_x, a_y, a_z\right)$. This is usually called a *position vector* or a *displacement vector*. The scalars a_x, a_y, and a_z are termed the *scalar components* of $\mathbf{a}$, and $\mathbf{a}$ is said to be fully characterized by a specification of this ordered list of three components:

$$\mathbf{a} = \left(a_x, a_y, a_z\right). \tag{A.1}$$

If the components in Eq. (A.1) are all real, $\mathbf{a}$ is said to be a *real vector*; if the components are complex numbers, $\mathbf{a}$ is called a *complex vector*. A *null vector* $\mathbf{0}$ is a vector with all components equal to zero. If all the components of a vector $\mathbf{a}$ are equal to one, $\mathbf{a}$ is called a *unitary vector*.

Two vectors, $\mathbf{a}$ and $\mathbf{b}$, are said to be equal when all their components are identical: $a_x = b_x$, $a_y = b_y$, and $a_z = b_z$. As a result, the vectorial equation $\mathbf{a} = \mathbf{b}$ corresponds to three scalar equations.

Now let us draw three *unit vectors*, $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$, of unit length (i.e., $|\mathbf{e}_x| = \left|\mathbf{e}_y\right| = |\mathbf{e}_z| = 1$) along the positive x-, y-, and z-axes. Since the scalar components a_x, a_y, and a_z of $\mathbf{a}$ are obtained by a projection of $\mathbf{a}$ on the x-, y-, and z-axes (Figure A.1), each three-dimensional vector $\mathbf{a}$ can be represented as follows:

$$\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z, \tag{A.2}$$

where we followed the rules of vector addition and scalar multiplication, as defined in the following section (§A.3).

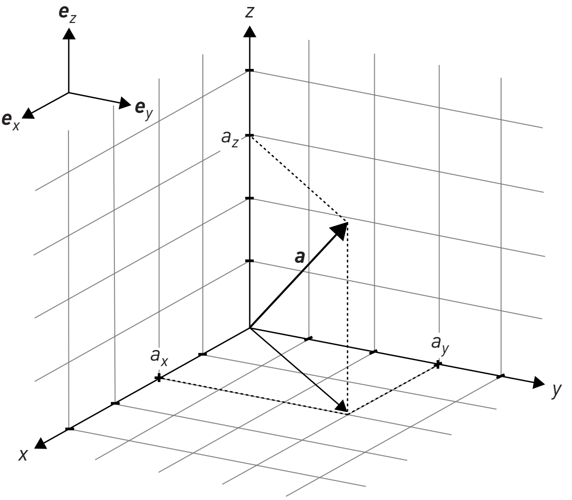
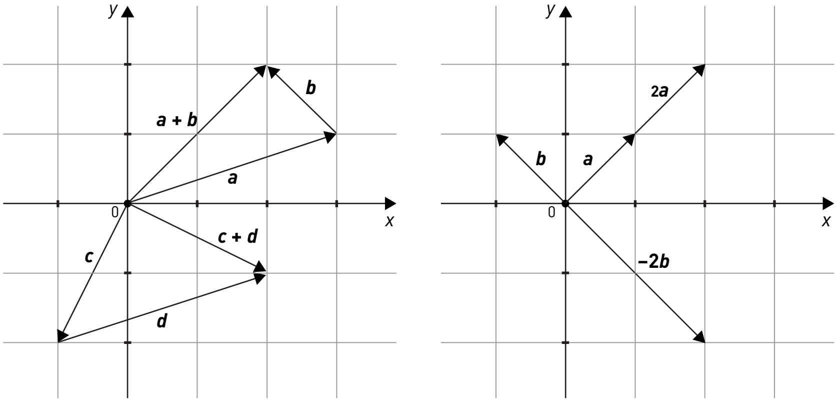


FIGURE A.1 Representation of a three-dimensional vector $\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z$ in a Cartesian $\Sigma\left(x, y, z\right)$ reference frame (with $a_x = 2$, $a_y = 3$, and $a_z = 3$).



Vector addition

Scalar multiplication

FIGURE A.2 Examples of *vector addition* and *scalar multiplication*. When written in terms of their components, we obtain $\mathbf{a} + \mathbf{b} = \left(3\mathbf{e}_x + 1\mathbf{e}_y\right) + \left(-1\mathbf{e}_x + 1\mathbf{e}_y\right) = 2\mathbf{e}_x + 2\mathbf{e}_y$ for the first addition, and $\mathbf{c} + \mathbf{d} = \left(-1\mathbf{e}_x - 2\mathbf{e}_y\right) + \left(3\mathbf{e}_x + 1\mathbf{e}_y\right) = 2\mathbf{e}_x - 1\mathbf{e}_y$ for the second. For scalar multiplication, this yields $2\mathbf{a} = 2\left(\mathbf{e}_x + \mathbf{e}_y\right) = 2\mathbf{e}_x + 2\mathbf{e}_y$ for the first multiplication, and $-2\mathbf{b} = -2\left(-\mathbf{e}_x + \mathbf{e}_y\right) = 2\mathbf{e}_x - 2\mathbf{e}_y$ for the second.

A.3 VECTOR OPERATIONS

The sum of two vectors **a** and **b** is obtained by sliding the first vector **a** in such a way that, while keeping its direction fixed, its head touches the tail of vector **b**. The sum $\mathbf{a} + \mathbf{b}$ is then defined as a new vector with an origin and an end point that coincide with the tail of **a** and the head of **b**, respectively. This is called *vector addition* (Figure A.2).

The product of a vector **a** with a scalar c results in a new vector $c\mathbf{a}$ with a length that is $|c|$ times the magnitude of **a** and with a direction that is the same as **a** when c is positive, or opposite to the direction of **a** when c is negative. We say that vector **a** has been rescaled by a factor of c. This is called *scalar multiplication* (Figure A.2).

The representation of a vector in terms of its components in Eq. (A.2) allows us to derive analytic expressions for both vector addition and scalar multiplication. The sum of two vectors **a** and **b** is thus

$$\begin{aligned}\mathbf{a}+\mathbf{b}&=\left(a_x\mathbf{e}_x+a_y\mathbf{e}_y+a_z\mathbf{e}_z\right)+\left(b_x\mathbf{e}_x+b_y\mathbf{e}_y+b_z\mathbf{e}_z\right)\\&=(a_x+b_x)\,\mathbf{e}_x+\left(a_y+b_y\right)\mathbf{e}_y+(a_z+b_z)\,\mathbf{e}_z.\end{aligned}\tag{A.3}$$

Vector addition is both *commutative* (i.e., $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$) and *associative* (i.e., $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$). The difference between **a** and **b**, called *vector subtraction*, is similarly defined:

$$\begin{aligned}\mathbf{a}-\mathbf{b}&=a_x\mathbf{e}_x+a_y\mathbf{e}_y+a_z\mathbf{e}_z-\left(b_x\mathbf{e}_x+b_y\mathbf{e}_y+b_z\mathbf{e}_z\right)\\&=(a_x-b_x)\,\mathbf{e}_x+\left(a_y-b_y\right)\mathbf{e}_y+(a_z-b_z)\,\mathbf{e}_z.\end{aligned}\tag{A.4}$$

Finally, a scalar multiplication of **a** by c results in the new vector

$$c\mathbf{a}=c\left(a_x\mathbf{e}_x+a_y\mathbf{e}_y+a_z\mathbf{e}_z\right)=ca_x\mathbf{e}_x+ca_y\mathbf{e}_y+ca_z\mathbf{e}_z.\tag{A.5}$$

When $c=0$, the null vector **0** is obtained; if $c=-1$, the vector **a** is said to be *reflected* about the origin and its direction is reversed: $(-1)\mathbf{a}=-\mathbf{a}$.

A.4 THE INNER PRODUCT

The *dot product* or *inner product* of two vectors, **a** and **b**, is defined as

$$\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\,|\mathbf{b}|\cos\phi,\tag{A.6}$$

where ϕ is the angle between **a** and **b**, and where $|\mathbf{a}|$ and $|\mathbf{b}|$ are the magnitude of **a** and **b**. Since $|\mathbf{a}|$, $|\mathbf{b}|$, and $\cos\phi$ are all scalars, the result of $\mathbf{a}\cdot\mathbf{b}$ is also a scalar—hence the alternative name *scalar product*. Note also that $\mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{a}$; the dot product is said to be *commutative*.[1] Based on this definition, the scalar product of a vector **a** with itself gives the squared magnitude of **a**:

$$\mathbf{a}\cdot\mathbf{a}=|\mathbf{a}|\,|\mathbf{a}|\cos 0=|\mathbf{a}|^2=\|\mathbf{a}\|.\tag{A.7}$$

The squared magnitude of **a** is also called the *Euclidean norm* of **a**, denoted by the symbol $\|\mathbf{a}\|$. Two vectors, **a** and **b**, are said to be *orthogonal* when their inner product equals zero—that is, when they form a right angle:

$$\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\,|\mathbf{b}|\cos\pi/2=0.\tag{A.8}$$

When two vectors, **a** and **b**, are orthogonal *and* of unit length ($|\mathbf{a}|=|\mathbf{b}|=1$), then **a** and **b** are said to be *orthonormal*. Since the unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are mutually perpendicular and all of length 1, their inner products equal

$$\mathbf{e}_x\cdot\mathbf{e}_x=\mathbf{e}_y\cdot\mathbf{e}_y=\mathbf{e}_z\cdot\mathbf{e}_z=1;\tag{A.9}$$

$$\mathbf{e}_x\cdot\mathbf{e}_y=\mathbf{e}_x\cdot\mathbf{e}_z=\mathbf{e}_y\cdot\mathbf{e}_z=0.\tag{A.10}$$

The set of pairwise orthogonal vectors in Eq. (A.10) is said to form an *orthogonal set*; this is denoted as $\mathbf{e}_x\perp\mathbf{e}_y\perp\mathbf{e}_z$. Since all vectors in this set are of unit length, Eq. (A.9) holds and the set is said to be *orthonormal*.

Following Eqs. (A.9) and (A.10), the dot product in Eq. (A.6) can be rewritten in terms of the components of **a** and **b**:

$$\begin{aligned}\mathbf{a}\cdot\mathbf{b}&=\left(a_x\mathbf{e}_x+a_y\mathbf{e}_y+a_z\mathbf{e}_z\right)\cdot\left(b_x\mathbf{e}_x+b_y\mathbf{e}_y+b_z\mathbf{e}_z\right)\\&=a_xb_x\mathbf{e}_x\cdot\mathbf{e}_x+a_yb_y\mathbf{e}_y\cdot\mathbf{e}_y+a_zb_z\mathbf{e}_z\cdot\mathbf{e}_z\\&=a_xb_x+a_yb_y+a_zb_z.\end{aligned}\tag{A.11}$$

---

[1] This applies to real vectors only. For complex vectors, we have $\mathbf{a}\cdot\mathbf{b}=\sum_i a_i^* b_i=(\mathbf{b}\cdot\mathbf{a})^*$.

In complete analogy, the dot product in Eq. (A.7) of vector **a** with itself equals

$$\begin{aligned}\mathbf{a}\cdot\mathbf{a} &= \left(a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z\right)\cdot\left(a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z\right)\\ &= a_x^2 + a_y^2 + a_z^2 = |\mathbf{a}|^2 = \|\mathbf{a}\|.\end{aligned} \tag{A.12}$$

The magnitude of **a** is therefore given by

$$|\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}, \tag{A.13}$$

which is just an extension of the Pythagorean theorem for the three sides of a right-angle triangle. A vector **a** with length $|\mathbf{a}| = 1$ is said to be *normalized*.

A.5 THE OUTER PRODUCT

Another kind of vector product, denoted $\mathbf{a}\times\mathbf{b}$ and termed the *cross product* or *outer product*, has to be defined within three-dimensional Euclidean space $\mathbb{R}^3$. The vector product of two vectors, **a** and **b**, yields another vector $\mathbf{a}\times\mathbf{b}$ (rather than a scalar $\mathbf{a}\cdot\mathbf{b}$), with magnitude

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\sin\phi, \tag{A.14}$$

where ϕ is the angle between the vectors **a** and **b**. The vector $\mathbf{a}\times\mathbf{b}$ is perpendicular to the plane formed by the vectors **a** and **b**, and its direction is given by the *right-hand rule* (Figure A.3) to form a right-handed system with **a** and **b** (just as the x-, y-, and z-axes also form a right-handed system). As a result, the cross product is *anticommutative*:

$$\mathbf{a}\times\mathbf{b} = -\left(\mathbf{b}\times\mathbf{a}\right). \tag{A.15}$$

Based on this definition, the cross products of the unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ can be easily calculated:

$$\begin{array}{lll}\mathbf{e}_x\times\mathbf{e}_x = 0; & \mathbf{e}_y\times\mathbf{e}_y = 0; & \mathbf{e}_z\times\mathbf{e}_z = 0;\\ \mathbf{e}_x\times\mathbf{e}_y = \mathbf{e}_z; & \mathbf{e}_y\times\mathbf{e}_z = \mathbf{e}_x; & \mathbf{e}_z\times\mathbf{e}_x = \mathbf{e}_y;\\ \mathbf{e}_y\times\mathbf{e}_x = -\mathbf{e}_z; & \mathbf{e}_z\times\mathbf{e}_y = -\mathbf{e}_x; & \mathbf{e}_x\times\mathbf{e}_z = -\mathbf{e}_y.\end{array} \tag{A.16}$$

According to this set of equations, the outer product of **a** and **b** can be written in terms of its scalar components:

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= \left(a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z\right)\times\left(b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z\right)\\ &= a_xb_y\mathbf{e}_x\times\mathbf{e}_y + a_xb_z\mathbf{e}_x\times\mathbf{e}_z + a_yb_x\mathbf{e}_y\times\mathbf{e}_x\\ &\qquad + a_yb_z\mathbf{e}_y\times\mathbf{e}_z + a_zb_x\mathbf{e}_z\times\mathbf{e}_x + a_zb_y\mathbf{e}_z\times\mathbf{e}_y\\ &= a_xb_y\mathbf{e}_z - a_xb_z\mathbf{e}_y - a_yb_x\mathbf{e}_z + a_yb_z\mathbf{e}_x + a_zb_x\mathbf{e}_y - a_zb_y\mathbf{e}_x\\ &= \left(a_yb_z - a_zb_y\right)\mathbf{e}_x + \left(a_zb_x - a_xb_z\right)\mathbf{e}_y + \left(a_xb_y - a_yb_x\right)\mathbf{e}_z.\end{aligned} \tag{A.17}$$

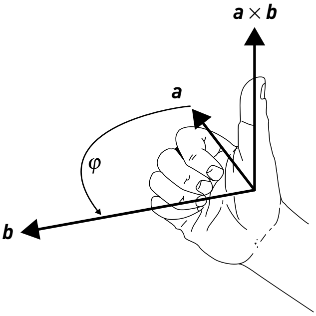


FIGURE A.3 The *right-hand rule*. Curl your fingers in such a way that it represents the rotation of vector **a** to vector **b**. Your thumb then points in the direction of the vector product $\mathbf{a}\times\mathbf{b}$, which is perpendicular to the plane formed by the vectors **a** and **b**. Notice that the vector product $\mathbf{b}\times\mathbf{a}$ points in the opposite direction.

To avoid making this cumbersome calculation over and over again, a useful mnemonic consists of representing the cross product $\mathbf{a} \times \mathbf{b}$ as a 3×3 determinant and expanding it as follows:

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} \\ &= \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \mathbf{e}_x - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} \mathbf{e}_y + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \mathbf{e}_z \\ &= \left(a_y b_z - a_z b_y\right) \mathbf{e}_x + \left(a_z b_x - a_x b_z\right) \mathbf{e}_y + \left(a_x b_y - a_y b_x\right) \mathbf{e}_z. \end{aligned} \tag{A.18}$$

A.6 HIGHER DIMENSIONAL VECTORS

So far, the discussion of vector algebra has been restricted to three-dimensional real vectors. Within the realm of abstract mathematics, however, there is no intrinsic difficulty in working with higher dimensional vectors. Therefore, let us start by generalizing Eq. (A.2) to vectors in n-dimensional Euclidean space $\mathbb{R}^n$ (also called *hyperspace*). Denoting the scalar components of an n-dimensional vector $\mathbf{a}$ by $a_1, a_2, \ldots, a_n$ and the n unit vectors by the symbols $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$,

$$\mathbf{a} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \ldots + a_n \mathbf{e}_n. \tag{A.19}$$

In summation notation,

$$\mathbf{a} = \sum_{k=1}^{n} a_k \mathbf{e}_k. \tag{A.20}$$

As shorthand for this last expression, use can be made of Einstein's *summation convention*, according to which Eq. (A.20) is written simply as $a_k \mathbf{e}_k$ (without the summation symbol $\sum$); repeated indices imply summation over all possible values of these indices (in this case, k). This notation was introduced by Albert Einstein (1879–1955) in 1916 in his ground breaking publication *The Foundation of the General Theory of Relativity*.[2]

Furthermore, we refer to the number n in Eq. (A.20) as the *order* of $\mathbf{a}$; vectors of order 1 are called *scalars*. For reasons of compactness, vectors of order n are also termed n-vectors; a vector in three-dimensional space, for instance, is called a 3-vector.

Two vectors, $\mathbf{a}$ and $\mathbf{b}$, of order n are said to be equal (i.e., $\mathbf{a} = \mathbf{b}$) when their corresponding scalar components are equal: $a_1 = b_1$, $a_2 = b_2$, …, $a_n = b_n$. One vector equation is therefore equivalent to n scalar equations.

The sum of two n-dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= (a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \ldots + a_n \mathbf{e}_n) + (b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + \ldots + b_n \mathbf{e}_n) \\ &= (a_1 + b_1)\,\mathbf{e}_1 + (a_2 + b_2)\,\mathbf{e}_2 + \ldots + (a_n + b_n)\,\mathbf{e}_n \\ &= \sum_{k=1}^{n} (a_k + b_k)\mathbf{e}_k. \end{aligned} \tag{A.21}$$

The difference of $\mathbf{a}$ and $\mathbf{b}$ is defined along the same lines. Also, the product of an n-dimensional vector $\mathbf{a}$ with a scalar s gives

$$s\mathbf{a} = s\,(a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \ldots + a_n \mathbf{e}_n) = \sum_{k=1}^{n} s a_k \mathbf{e}_k. \tag{A.22}$$

---

[2] A. Einstein. “The Foundation of the General Theory of Relativity. [Grundlage der allgemeinen Relativitätstheorie].” *Annalen der Physik* IV.49 (1916), pp. 769–822.

Generalizing from Eq. (A.11), the inner product $\mathbf{a} \cdot \mathbf{b}$ of two n-dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\mathbf{a} \cdot \mathbf{b} = \sum_{k=1}^{n} a_k b_k. \tag{A.23}$$

Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be *orthogonal* when $\mathbf{a} \cdot \mathbf{b} = 0$. Analogous to Eq. (A.13), the magnitude $|\mathbf{a}|$ of an n-dimensional vector $\mathbf{a}$ is defined as

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \left(a_1^2 + a_2^2 + \ldots + a_n^2\right)^{1/2} = \left[\sum_{i=1}^{n} a_i^2\right]^{1/2}. \tag{A.24}$$

Appendix B
Matrix algebra

B.1 CONCEPTUAL DEFINITION

In mathematics, a *matrix* is defined as a rectangular array of (real or complex) numbers. All matrices in this book are denoted by uppercase blackboard bold symbols, such as $\mathbb{A}$. Let $\mathbb{A}$ be a matrix consisting of m rows and n columns:

$$\mathbb{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}. \tag{B.1}$$

$\mathbb{A}$ is said to be an $m \times n$ (read: m by n) matrix; m and n are called the *dimensions* of $\mathbb{A}$. It is convenient in some cases to list the dimensions of a matrix underneath the matrix symbol; the matrix in Eq. (B.1) is then denoted as $\underset{m\times n}{\mathbb{A}}$.

Each entry (number) in a matrix is termed a *matrix element* and is denoted by the general symbol a_{ij}, with the subscripts $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, referring to the row and the column of that particular element. For an $m \times n$ matrix $\mathbb{A}$, there are, in total, mn matrix elements enclosed in square brackets.

B.2 SPECIAL MATRICES

Analogous to vectors, the matrix $\mathbb{A}$ in Eq. (B.1) is called a *real matrix* if all its matrix elements are real. If, on the other hand, the matrix elements are complex numbers, then $\mathbb{A}$ is said to be a *complex matrix* (see §B.6).

A matrix of dimension $1 \times n$ is called a *row matrix*; similarly, a matrix of dimension $m \times 1$ is termed a *column matrix*. Matrices with an equal number of rows and columns ($n = m$) are called *square matrices* of *order n*. The n components a_{ii} of a square matrix run from the upper left to the bottom right and are said to form the *main diagonal* of that matrix. The *cross diagonal* runs in the opposite direction, from the bottom left to the upper right:

$$\underset{1\times n \text{ row matrix}}{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}} \qquad \underset{m\times 1 \text{ column matrix}}{\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}} \qquad \underset{n\times n \text{ square matrix}}{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}$$

Square matrices for which $a_{ij} = a_{ji} \quad \forall i, j = 1, 2, \ldots, n$ are symmetric about the main diagonal and are therefore termed *symmetric matrices*. Square matrices for which $a_{ij} = -a_{ji}$

$\forall i, j = 1, 2, \ldots, n$ are said to be *antisymmetric* or *skew-symmetric*. Based on this definition, it follows that the diagonal elements a_{ii} of a skew-symmetric matrix have to be zero since $a_{ii} = -a_{ii}$.

Matrices with elements that are all zero are called *null matrices* and are denoted by the symbol $\mathbb{O}$. The *unit matrix*, or *identity matrix* $\mathbb{I}$, is a square matrix of order n with main diagonal elements that are all 1 and with off-diagonal elements equal to 0. When all the off-diagonal elements of a square matrix $\mathbb{A}$ of order n are 0, then $\mathbb{A}$ is called a *diagonal matrix*.

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Null matrix $\mathbb{O}$ Identity matrix $\mathbb{I}$ Diagonal matrix $\mathbb{A}$

B.3 MATRIX OPERATIONS

B.3.1 Equality

Two matrices, $\mathbb{A}$ and $\mathbb{B}$, are said to be *equal* when they have the same dimensions—say, m rows by n columns—and when each matrix element a_{ij} of $\mathbb{A}$ equals the corresponding element b_{ij} of $\mathbb{B}$: $a_{ij} = b_{ij}\ \forall i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n$. The matrix equation $\mathbb{A} = \mathbb{B}$ is therefore equivalent to mn scalar equations. Note that matrices of different dimensions cannot be compared.

B.3.2 Matrix transposition

Let us proceed by considering a number of basic matrix manipulations called *matrix transposition*, *matrix addition*, *matrix subtraction*, and *scalar multiplication*. The *transpose* of an $m \times n$ matrix $\mathbb{A}$ is defined to be an $n \times m$ matrix, written $\mathbb{A}^{\mathrm{T}}$, and is obtained by changing all rows (columns) of $\mathbb{A}$ into columns (rows) of $\mathbb{A}^{\mathrm{T}}$. Denoting the matrix elements of $\mathbb{A}^{\mathrm{T}}$ by a^{T}_{ij}, we obtain the following relationship between the elements of $\mathbb{A}$ and $\mathbb{A}^{\mathrm{T}}$:

$$a^{\mathrm{T}}_{ij} = a_{ji} \qquad \forall i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n. \tag{B.2}$$

For example, transposing a 4 × 3 matrix $\mathbb{A}$, results in a 3 × 4 matrix $\mathbb{A}^{\mathrm{T}}$:

$$\mathbb{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{bmatrix};\ \mathbb{A}^{\mathrm{T}} = \begin{bmatrix} a & d & g & j \\ b & e & h & k \\ c & f & i & l \end{bmatrix}. \tag{B.3}$$

Clearly, when a transposed matrix $\mathbb{A}^{\mathrm{T}}$ is transposed a second time, the initial matrix $\mathbb{A}$ is reobtained—that is, $\left(\mathbb{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathbb{A}$. The transpose of a square matrix of order n is another square matrix of order n. Transposition of a symmetric matrix results in the original matrix (i.e., $\mathbb{A}^{\mathrm{T}} = \mathbb{A}$); the transpose of a skew-symmetric matrix is equal to this matrix multiplied by minus one (i.e., $\mathbb{A}^{\mathrm{T}} = -\mathbb{A}$).

B.3.3 Matrix addition and subtraction

The sum of two $m \times n$ matrices, $\mathbb{A}$ and $\mathbb{B}$, is obtained by adding the elements of $\mathbb{A}$ to the corresponding elements of $\mathbb{B}$:

$$\mathbb{A}+\mathbb{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}. \tag{B.4}$$

The result is a new $m \times n$ matrix, denoted $\mathbb{A}+\mathbb{B}$. It follows naturally that *matrix addition* is defined exclusively for matrices of the same dimension. Now suppose $\mathbb{C} = \mathbb{A}+\mathbb{B}$; the previous equation can then be rewritten in terms of the matrix elements of $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$:

$$c_{ij} = a_{ij} + b_{ij} \qquad \forall i = 1,2,\ldots,m;\ j = 1,2,\ldots,n. \tag{B.5}$$

Matrix addition is said to be *commutative* (i.e., $\mathbb{A}+\mathbb{B} = \mathbb{B}+\mathbb{A}$) since $a_{ij}+b_{ij} = b_{ij}+a_{ij}$. Moreover, matrix addition is *associative*: $(\mathbb{A}+\mathbb{B})+\mathbb{C} = \mathbb{A}+(\mathbb{B}+\mathbb{C})$. Two matrices can also be *subtracted*. That is, if $\mathbb{C} = \mathbb{A}-\mathbb{B}$, then

$$c_{ij} = a_{ij} - b_{ij} \qquad \forall i = 1,2,\ldots,m;\ j = 1,2,\ldots,n. \tag{B.6}$$

B.3.4 Scalar multiplication

The *scalar multiplication* of an $m \times n$ matrix $\mathbb{A}$ with a scalar c is obtained by multiplying *each* entry a_{ij} of $\mathbb{A}$ by c:

$$c\mathbb{A} = c \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}. \tag{B.7}$$

In other words, if $\mathbb{D} = c\mathbb{A}$, then $d_{ij} = ca_{ij}\ \forall i = 1,2,\ldots,m;\ j = 1,2,\ldots,n$. Multiplying matrix $\mathbb{A}$ with c is often referred to as *scaling*, analogous to the scalar multiplication of vectors. The null matrix $\mathbb{O}$ is obtained when $\mathbb{A}$ is multiplied by zero.

B.4 MATRIX PRODUCTS

B.4.1 Definition

A pair of matrices $\mathbb{A}$ and $\mathbb{B}$ can also be multiplied, provided that the number of columns in $\mathbb{A}$ is equal to the number of rows in $\mathbb{B}$. When this is the case, the matrices $\mathbb{A}$ and $\mathbb{B}$ are said to be *product conforming*. That is to say, if $\mathbb{A}$ is a matrix of dimension $m \times n$ and $\mathbb{B}$ is of dimension $n \times p$, then $\mathbb{A}$ and $\mathbb{B}$ conform and *matrix multiplication* is defined. The resulting matrix product, denoted[1]

$$\underset{m\times n}{\mathbb{A}} \times \underset{n\times p}{\mathbb{B}} = \underset{m\times p}{\mathbb{C}}, \tag{B.8}$$

---

[1] For notational simplicity, the explicit multiplication sign is omitted most of the time. That is, an expression with consecutive matrices is automatically understood to be a matrix product.

is an $m \times p$ matrix with elements given by the following formula:

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj}, \qquad\qquad \forall i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, p. \end{aligned} \tag{B.9}$$

The matrix element c_{ij} is thus obtained by multiplying each element a_{ik} of the ith row of $\mathbb{A}$ with the corresponding element b_{kj} of the jth column of $\mathbb{B}$ and then adding these n products. In matrix notation,

$$\underset{m \times n \text{ matrix } \mathbb{A}}{\begin{bmatrix} \cdot & \cdot & \cdots & \cdot \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot \end{bmatrix}} \underset{n \times p \text{ matrix } \mathbb{B}}{\begin{bmatrix} \cdot & b_{1j} & \cdots & \cdot \\ \cdot & b_{2j} & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & b_{nj} & \cdots & \cdot \end{bmatrix}} = \underset{m \times p \text{ matrix } \mathbb{C}}{\begin{bmatrix} \cdot & \cdot & \cdots & \cdot \\ \cdot & c_{ij} & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot \end{bmatrix}}. \tag{B.10}$$

Alternatively, the previous matrix product can also be represented with the help of *Falk's scheme*, which is constructed as follows:

$$\begin{array}{cc} & \begin{bmatrix} \cdot & b_{1j} & \cdots & \cdot \\ \cdot & b_{2j} & \cdots & \cdot \\ \vdots & \downarrow & \ddots & \vdots \\ \cdot & b_{nj} & \cdots & \cdot \end{bmatrix} \\ \begin{bmatrix} \cdot & \cdot & \cdots & \cdot \\ a_{i1} & a_{i2} & \rightarrow & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot \end{bmatrix} & \begin{bmatrix} \cdot & \cdot & \cdots & \cdot \\ \cdot & c_{ij} & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot \end{bmatrix} \end{array}. \tag{B.11}$$

B.4.2 Properties of matrix products

Matrices are highly relevant for group theory. The reason is that their products obey some simple properties, which represent the abstract combination rules in groups.

1. *Commutativity*. Although the product $\underset{m\times n}{\mathbb{A}} \times \underset{n\times p}{\mathbb{B}}$ in Eq. (B.8) is warranted, the reverse product $\underset{n\times p}{\mathbb{B}} \times \underset{m\times n}{\mathbb{A}}$ is not necessarily defined since p might be different from m. Even when $p = m$, the product $\mathbb{B}\mathbb{A}$ is usually different from $\mathbb{A}\mathbb{B}$. That is, matrix multiplication is, in general, not *commutative*—in other words, $\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$. If $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A}$, then $\mathbb{A}$ and $\mathbb{B}$ are said to *commute*; diagonal matrices always commute.
2. *Associativity*. However, matrix multiplication obeys the rules of *associativity*:

$$(\mathbb{A}\mathbb{B})\,\mathbb{C} = \mathbb{A}\mathbb{B}\mathbb{C} = \mathbb{A}\,(\mathbb{B}\mathbb{C}). \tag{B.12}$$

3. *Distributivity*. Both left and right *distributivity* are also seen to hold true:

$$\begin{aligned} \mathbb{A}\,(\mathbb{B} + \mathbb{C}) &= \mathbb{A}\mathbb{B} + \mathbb{A}\mathbb{C}; \\ (\mathbb{A} + \mathbb{B})\,\mathbb{C} &= \mathbb{A}\mathbb{C} + \mathbb{B}\mathbb{C}. \end{aligned} \tag{B.13}$$

B.4.3 The inner product in matrix notation

An important connection can be made at this point with the inner product of vectors, defined in §A.4. Since an n-dimensional vector $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ is completely specified by its scalar components $a_1, a_2, \ldots, a_n$, a vector **a** can be represented as a $1 \times n$ row matrix or as an $n \times 1$ column matrix with matrix elements that correspond to the n components of **a**. Row and column matrices are therefore synonymously termed *row* and *column vectors.* Both ways of representing a vector are used in this book. Coordinate vectors are written as column matrices; function spaces are written as row matrices.

Transposing an $n \times 1$ column vector gives a $1 \times n$ row vector, and vice versa. Now, let **a** be an n-dimensional vector represented by an $n \times 1$ column matrix. We may then associate with this column vector a corresponding *dual vector*, denoted $\mathbf{a}^{\mathrm{T}}$, which is defined as the transposed form, or row form, of **a** and which has dimensions of $1 \times n$:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}; \ \mathbf{a}^{\mathrm{T}} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}. \tag{B.14}$$

According to Eq. (B.9), the matrix product of **a** and **b** is given by

$$\mathbf{a}^{\mathrm{T}}\mathbf{b} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{k=1}^{n} a_k b_k = \mathbf{a} \cdot \mathbf{b}. \tag{B.15}$$

This is nothing else than the dot product of **a** and **b**. It follows that, in general, the matrix product $\mathbb{A}\mathbb{B} = \mathbb{C}$ in Eq. (B.10) is equivalent to mp dot products. That is, each element c_{ij} in $\mathbb{C}$ is obtained by making the inner product of the ith row vector of $\mathbb{A}$ with the jth column vector of $\mathbb{B}$.

B.4.4 The transpose of a matrix product

An important property of matrix multiplication is that the transpose of a matrix product is equal to the product of the transposes written in *reverse* order:

$$(\mathbb{A}\mathbb{B})^{\mathrm{T}} = \mathbb{B}^{\mathrm{T}}\mathbb{A}^{\mathrm{T}}. \tag{B.16}$$

This statement may look counterintuitive at first; many of you may, quite naturally, expect $(\mathbb{A}\mathbb{B})^{\mathrm{T}}$ to equal $\mathbb{A}^{\mathrm{T}}\mathbb{B}^{\mathrm{T}}$ instead. There is, however, a logical explanation for the reversal in Eq. (B.16). Let $\mathbb{A}$ be an $m \times n$ matrix and $\mathbb{B}$ be an $n \times p$ matrix. Denoting the result of the matrix product $\mathbb{A}\mathbb{B}$ by $\mathbb{C}$ gives

$$\underset{m\times n}{\mathbb{A}}\ \underset{n\times p}{\mathbb{B}} = \underset{m\times p}{\mathbb{C}}. \tag{B.17}$$

The transpose of $\mathbb{A}\mathbb{B}$ is then $(\mathbb{A}\mathbb{B})^{\mathrm{T}} = \mathbb{C}^{\mathrm{T}}$, with $\mathbb{C}^{\mathrm{T}}$ a $p \times m$ matrix. Following Eq. (B.16) leads to the equation

$$\underset{p\times n}{\mathbb{B}^{\mathrm{T}}}\ \underset{n\times m}{\mathbb{A}^{\mathrm{T}}} = \underset{p\times m}{\mathbb{C}^{\mathrm{T}}}, \tag{B.18}$$

which clearly holds true. The intuitively more natural (but erroneous) option, according to which $(\mathbb{A}\mathbb{B})^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}}\mathbb{B}^{\mathrm{T}}$ yields

$$\underset{n\times m}{\mathbb{A}^{\mathrm{T}}}\ \underset{p\times n}{\mathbb{B}^{\mathrm{T}}} = \underset{p\times m}{\mathbb{C}^{\mathrm{T}}}, \tag{B.19}$$

which does not hold because m is usually different from p, and matrix multiplication is therefore not defined for Eq. (B.19). Even if $m = p$, the result of $\mathbb{A}^{\mathrm{T}}\mathbb{B}^{\mathrm{T}}$ is an $n \times n$ matrix, which is different from the $p \times m$ matrix $\mathbb{C}^{\mathrm{T}}$.

B.5 TRACE AND DETERMINANT OF A SQUARE MATRIX

The algebra of square matrices is controlled by two important characteristics: the *trace* and the *determinant*. These respectively conserve the sum and product operations.

B.5.1 The trace of a square matrix

The trace of a square matrix, denoted Tr, is defined as the sum of the diagonal elements. So, for $\mathbb{A}$ being an $n \times n$ square matrix, we have

$$\operatorname{Tr}(\mathbb{A}) = \sum_{i=1}^{n} a_{ii}. \tag{B.20}$$

The trace of a sum of matrices is equal to the sum of their traces. In other words, the trace conserves matrix summation:

$$\operatorname{Tr}(\mathbb{A}+\mathbb{B}) = \sum_{i=1}^{n} (a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \operatorname{Tr}(\mathbb{A}) + \operatorname{Tr}(\mathbb{B}). \tag{B.21}$$

We can also verify that the trace is immune for matrix transposition, since transposition leaves the matrix diagonal unchanged:

$$\operatorname{Tr}\left(\mathbb{A}^{\mathrm{T}}\right) = \operatorname{Tr}(\mathbb{A}). \tag{B.22}$$

B.5.2 The determinant of a square matrix

The determinant of a square matrix is perhaps one of the most important matrix characteristics, because it is an indispensable tool to solve systems of linear equations. Its definition relies on the concept of permutations. The permutation of an ordered set is an operation that only rearranges the order in this set. Hence, a permutation is a one-to-one mapping of a given set onto itself. So, a permutation σ on the set $(1,2,3,\ldots,n)$ may be defined as

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma_1 & \sigma_2 & \ldots & \sigma_n \end{pmatrix}, \tag{B.23}$$

where the $\{\sigma_i\}$ set is simply a rearrangement of the numbers $1,2,\ldots,n$. As an example, in Chapter 5, we make use of a cyclic permutation of the indices x,y,z. For this set of three elements, two cyclic permutations are possible:

$$\pi_1 = \begin{pmatrix} x & y & z \\ y & z & x \end{pmatrix}, \tag{B.24}$$

$$\pi_2 = \begin{pmatrix} x & y & z \\ z & x & y \end{pmatrix}. \tag{B.25}$$

In π_1, x becomes y, y becomes z, and z is replaced with x, thereby closing the cycle. In π_2, the cycle is traversed in the opposite sense: z goes back to y, y goes back to x, and x goes back to z. Hence, both operations are the inverse of each other.

In general, to every permutation there corresponds an inverse permutation that puts the elements back in the starting order. There is also a unit permutation that simply maps each element onto itself. Applying two permutations in succession is again a permutation, and the combination of permutations is associative; in short, permutations form a group that is called the *symmetric group*, denoted S_n. This group is discrete of order $n!$ Only $n-1$ of its elements are cyclic permutations (or n if you include the unit permutation).

A permutation that interchanges two elements only is called a *transposition.* As an example, the transposition of the elements i and j is given by

$$P_{ij} = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ 1 & 2 & \dots & j & \dots & i & \dots & n \end{pmatrix}. \tag{B.26}$$

Every permutation can be obtained as the result of a sequence of transpositions. As an example, to change the sequence $(1,2,3,4)$ to $(2,3,4,1)$, we need at least three transpositions:

$$(1,2,3,4) \xrightarrow{P_{12}} (2,1,3,4) \xrightarrow{P_{13}} (2,3,1,4) \xrightarrow{P_{14}} (2,3,4,1). \tag{B.27}$$

The same result may be realized by other transposition pathways; but, in any case, we always need an odd number of transpositions to achieve the overall permutation. Hence, a permutation can be characterized by a parity or sign that refers to the parity of the number of transpositions in the overall permutation. This parity is denoted as $\mathrm{sgn}(\sigma)$, and is defined as follows:

$$\begin{aligned} &\text{Parity of } \sigma \text{ even: } \mathrm{sgn}(\sigma) = +1; \\ &\text{parity of } \sigma \text{ odd: } \mathrm{sgn}(\sigma) = -1. \end{aligned} \tag{B.28}$$

The definition of the determinant of an $n \times n$ square matrix $\mathbb{A}$ is based on the permutations of the n indices $(1,2,\dots,n)$. The determinant is denoted as $\det(\mathbb{A})$ or as $|\mathbb{A}|$ and is given by

$$\det(\mathbb{A}) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma_1} a_{2\sigma_2} a_{3\sigma_3} \dots a_{n\sigma_n}. \tag{B.29}$$

The terms in this sum may be visualized as paths through the matrix, starting at some position in the top row and going down gradually, row by row, to the bottom row, thereby making sure that every column is visited only once. There are exactly $n!$ of such paths. For a 2×2 matrix there are only two paths: one along the diagonal and the other, with opposite sign, along the cross diagonal:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{B.30}$$

For a 3×3 matrix, a convenient expression is offered by the *minor* expansion. The minor of a given element a_{ij} is equal to the determinant of the submatrix obtained by deleting the ith row and the jth column. The *cofactor* is the "signed" minor, multiplied by $(-1)^{i+j}$. The determinant is then equal to the sum of the products obtained by multiplying the elements of the first row by their respective cofactors:

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \\ &\quad + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} \\ &\quad + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \tag{B.31}$$

This expansion appears in Appendix A to express the outer product of two vectors (see Eq. A.18). It is a particular case of the Laplace expansion formula for determinants. A more general form of this formula is discussed in Appendix E.

An important property of the determinant is that it conserves the matrix product. Hence, the determinant of a product of matrices is equal to the product of the respective determinants:

$$\det(\mathbb{A} \times \mathbb{B}) = \det(\mathbb{A})\det(\mathbb{B}). \tag{B.32}$$

The determinant literally "determines" whether a matrix has an inverse. In fact, for a matrix to be invertible, it is sufficient that its determinant be nonzero. Matrices with this property are called *nonsingular*. Appendix E elaborates this point.

Finally, the determinant does not change under transposition:

$$\det\left(\mathbb{A}^{\mathrm{T}}\right) = \det(\mathbb{A}). \tag{B.33}$$

This property can be easily understood from the definition of the determinant. The role of the permutations of the column labels in this definition was to generate all the paths from the top row to the bottom row that were visiting each column only once. It is clear that this collection of paths will equally well contain all the paths that are traced when going from the column on the left to the column on the right, that visit every row only once. Hence, in the definition, the permutations can be applied equally well to the row indices:

$$\det(\mathbb{A}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma_1 1} a_{\sigma_2 2} a_{\sigma_3 3} \ldots a_{\sigma_n n}, \tag{B.34}$$

from which Eq. (B.29) follows.

B.6 COMPLEX MATRICES

A *complex matrix* $\mathbb{C}$ of dimension $m \times n$ is a matrix with complex matrix elements c_{ij} $(i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n)$. Since every complex number c can be put in the form $c = R + iI$, with R and I real numbers and $i = \sqrt{-1}$, the matrix $\mathbb{C}$ can be written as

$$\mathbb{C} = \begin{bmatrix} R_{11} + iI_{11} & R_{12} + iI_{12} & R_{13} + iI_{13} & \cdots & R_{1n} + iI_{1n} \\ R_{21} + iI_{21} & R_{22} + iI_{22} & R_{23} + iI_{23} & \cdots & R_{2n} + iI_{2n} \\ R_{31} + iI_{31} & R_{32} + iI_{32} & R_{33} + iI_{33} & \cdots & R_{3n} + iI_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{m1} + iI_{m1} & R_{m2} + iI_{m2} & R_{m3} + iI_{m3} & \cdots & R_{mn} + iI_{mn} \end{bmatrix}, \tag{B.35}$$

where R_{ij} and I_{ij} are the real and imaginary parts, respectively, of the matrix elements c_{ij}. The *complex conjugate* matrix $\mathbb{C}^*$ is obtained by changing the sign of all imaginary parts:

$$\mathbb{C}^* = \begin{bmatrix} R_{11} - iI_{11} & R_{12} - iI_{12} & R_{13} - iI_{13} & \cdots & R_{1n} - iI_{1n} \\ R_{21} - iI_{21} & R_{22} - iI_{22} & R_{23} - iI_{23} & \cdots & R_{2n} - iI_{2n} \\ R_{31} - iI_{31} & R_{32} - iI_{32} & R_{33} - iI_{33} & \cdots & R_{3n} - iI_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{m1} - iI_{m1} & R_{m2} - iI_{m2} & R_{m3} - iI_{m3} & \cdots & R_{mn} - iI_{mn} \end{bmatrix}. \tag{B.36}$$

In terms of its matrix elements,

$$c_{ij}^* = R_{ij} - iI_{ij} \qquad \forall i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n, \tag{B.37}$$

where c_{ij}^* denotes the complex conjugate of c_{ij}. Finally, the *conjugate transpose* (or *Hermitian ajdoint*, or *adjoint* for short) of $\mathbb{C}$, denoted $\mathbb{C}^\dagger$, is obtained by transposing the rows and columns of the conjugate matrix $\mathbb{C}^*$:

$$\mathbb{C}^\dagger = \left(\mathbb{C}^*\right)^{\mathrm{T}}. \tag{B.38}$$

In terms of its matrix elements,

$$\begin{aligned}(c_{ij})^{\dagger} &= c_{ji}^{*};\\ (R_{ij} + iI_{ij})^{\dagger} &= R_{ji} - iI_{ji}.\end{aligned} \tag{B.39}$$

B.7 ORTHOGONAL MATRICES

An *orthogonal matrix* $\mathbb{Q}$ is a square matrix of dimension $n \times n$, with real matrix elements q_{ij} $(i, j = 1, 2, \ldots, n)$, that satisfies the condition

$$\mathbb{Q}^{\mathrm{T}}\mathbb{Q} = \mathbb{Q}\mathbb{Q}^{\mathrm{T}} = \mathbb{I}, \tag{B.40}$$

where $\mathbb{I}$ is the $n \times n$ identity matrix and $\mathbb{Q}^{\mathrm{T}}$ is the transpose of $\mathbb{Q}$, as defined in §B.3.2. In other words, a real matrix $\mathbb{Q}$ is said to be orthogonal whenever its transpose is equal to its inverse:

$$\mathbb{Q}^{\mathrm{T}} = \mathbb{Q}^{-1}. \tag{B.41}$$

What is more, since the identity matrix has unit determinant, we obtain

$$\begin{aligned}\det\left(\mathbb{Q}^{\mathrm{T}}\mathbb{Q}\right) &= \det\mathbb{Q}^{\mathrm{T}}\det\mathbb{Q}\\ &= \det\mathbb{Q}\det\mathbb{Q}\\ &= (\det\mathbb{Q})^2\\ &= \det\mathbb{I} = 1,\end{aligned} \tag{B.42}$$

where we have used the matrix properties $\det(\mathbb{A}\mathbb{B}) = \det\mathbb{A}\det\mathbb{B}$ and $\det\mathbb{A}^{\mathrm{T}} = \det\mathbb{A}$. It follows that the determinant of an orthogonal matrix $\mathbb{Q}$ can only assume the values $+1$ or -1:

$$\det\mathbb{Q} = \pm 1. \tag{B.43}$$

B.8 UNITARY MATRICES

A *unitary matrix* $\mathbb{U}$ is a square matrix of dimension $n \times n$ with complex matrix elements $u_{ij} = R_{ij} + iI_{ij}$ $(i, j = 1, 2, \ldots, n)$ satisfying the condition

$$\mathbb{U}^{\dagger}\mathbb{U} = \mathbb{U}\mathbb{U}^{\dagger} = \mathbb{I}, \tag{B.44}$$

where $\mathbb{I}$ is the $n \times n$ identity matrix and $\mathbb{U}^{\dagger}$ is the conjugate transpose of $\mathbb{U}$, as defined in Eqs. (B.38) and (B.39). Note that if $\mathbb{U}$ is a matrix with only *real* matrix elements (i.e., $\mathbb{U}^* = \mathbb{U}$), then

$$\mathbb{U}^{\dagger} = \left(\mathbb{U}^*\right)^{\mathrm{T}} = \mathbb{U}^{\mathrm{T}}, \tag{B.45}$$

and Eq. (B.44) reduces to

$$\mathbb{U}^{\mathrm{T}}\mathbb{U} = \mathbb{U}\mathbb{U}^{\mathrm{T}} = \mathbb{I}. \tag{B.46}$$

That is, $\mathbb{U}$ is an *orthogonal* matrix (see §B.7). The set of unitary matrices thus contains all orthogonal matrices, but it is more general because matrices with complex elements are included as well. It follows from Eq. (B.44) that the conjugate transpose of a unitary matrix is equal to the inverse matrix:

$$\mathbb{U}^{\dagger} = \mathbb{U}^{-1}. \tag{B.47}$$

Also, since the identity matrix has unit determinant, we obtain

$$\begin{aligned}\det\left(\mathbb{U}^{\dagger}\mathbb{U}\right) &= \det\mathbb{U}^{\dagger}\det\mathbb{U}\\ &= \det\left(\mathbb{U}^*\right)^{\mathrm{T}}\det\mathbb{U}\\ &= \det\mathbb{U}^*\det\mathbb{U}\\ &= |\det\mathbb{U}|^2\\ &= \det\mathbb{I} = 1,\end{aligned} \tag{B.48}$$

where we have used the matrix properties $\det(\mathbb{A}\mathbb{B}) = \det\mathbb{A}\det\mathbb{B}$, and $\det\mathbb{A}^{\mathrm{T}} = \det\mathbb{A}$. It follows that for a unitary matrix with complex matrix elements, the determinant can be any complex number with *unit modulus* (i.e., a complex number with *absolute value* or *norm* 1):

$$|\det\mathbb{U}| = 1. \tag{B.49}$$

Matrices satisfying this condition are said to be *unimodular*.

Appendix C
Taylor and Maclaurin series

The process of expanding a function into a Taylor series was first devised by the English mathematician, Brook Taylor (1685–1731) in 1715.

Theorem C.1 (Taylor series): In mathematics, an arbitrary analytic function f can be represented as an infinite sum of terms calculated from the values of its derivatives at a single point a. This is called a *Taylor series*:

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} \left. \frac{d^n f(x)}{dx^n} \right|_{x=a}, \tag{C.1}$$

where $n!$ denotes the factorial of n and $\left. d^n f(x) \middle/ dx^n \right|_{x=a}$ denotes the nth derivative of f evaluated at point $x = a$. The zeroth derivative of f is defined to be f itself and $(x-a)^0$ and $0!$ are both defined to be one. ■

Corollary C.1 (Maclaurin series): If the Tailor series is centered at the origin ($a = 0$), the series is also termed a *Maclaurin series* after Scottish mathematician Colin Maclaurin (1698–1746). In sigma notation,

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \left. \frac{d^n f(x)}{dx^n} \right|_{x=0}, \tag{C.2}$$

where $n!$ is the factorial of n and $d^n f(x) / dx^n$ denotes the nth derivative of f evaluated at the origin $x = 0$. ■

Proof (Maclaurin series). Let us start by representing an arbitrary function f by a power series:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots. \tag{C.3}$$

Evaluating at $x = 0$, we have

$$f(0) = a_0. \tag{C.4}$$

Differentiating the function,

$$\frac{df(x)}{dx} = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \ldots. \tag{C.5}$$

Evaluating at $x = 0$,

$$\left. \frac{df(x)}{dx} \right|_{x=0} = a_1. \tag{C.6}$$

Differentiating the function again,

$$\frac{d^2 f(x)}{dx^2} = 2a_2 + 6a_3 x + 12a_4 x^2 + \ldots. \tag{C.7}$$

Evaluating at $x = 0$,

$$\frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=0} = a_2. \tag{C.8}$$

Generalizing,

$$a_n = \frac{1}{n!} \left. \frac{d^n f(x)}{dx^n} \right|_{x=0}, \tag{C.9}$$

where $d^n f(x) / dx^n$ is the nth derivative of $f(0)$. Substituting the respective values of a_n in the power expansion,

$$f(x) = f(0) + x \left. \frac{df(x)}{dx} \right|_{x=0} + \frac{x^2}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=0} + \frac{x^3}{3!} \left. \frac{d^3 f(x)}{dx^3} \right|_{x=0} + \ldots, \tag{C.10}$$

which is a particular case of the *Taylor series* (also known as a *Maclaurin series*), and which can be written in the more compact sigma notation:

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \left. \frac{d^n f(x)}{dx^n} \right|_{x=0}. \tag{C.11}$$

The zeroth derivative of f is defined to be f itself and x^0 and $0!$ are both defined to be one. This derivation can be easily generalized to prove the expression for a Taylor series as given in Eq. (C.1). ■

We have thus shown that any function f can be represented as an infinite sum of terms calculated from the values of its derivative at the origin. This is a fascinating conclusion to say the least, because it is sufficient to consider a *single point* of the function (in this case, the origin) to evaluate the function at *any* other point. As Richard Feynman poetically said during the Messenger Lectures of 1964: “Nature uses only the longest threads to weave her patterns, so each small piece of her fabric reveals the organization of the entire tapestry.”[1]

It must be noted, however, that a Taylor expansion still necessitates an *infinite* amount of knowledge about the function and its *infinite* number of derivatives at that particular point to represent the function *completely*. Of course, it is possible to end the summation at some point and to use only the first terms of the Taylor series, in which case the actual function f is *approximated* by the truncated Taylor series. Since the higher order terms become smaller and smaller as a result of the gradually increasing value of the factorial $n!$ and decreasing value of x^n (for $|x| < 1$), the approximation of f by using only a finite number of terms should be a reasonably good one. It is therefore useful to write the Taylor series in a slightly different format:

$$f(x) = f(0) + x \left. \frac{df(x)}{dx} \right|_{x=0} + \frac{x^2}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=0} + \mathcal{O}\left(\omega^3\right), \tag{C.12}$$

where we introduce the *order* of a Taylor series by the general symbol $\mathcal{O}(\omega^n)$, which indicates at which point the series is truncated. So, all the terms of order n or higher—denoted by the symbol $\mathcal{O}(\omega^n)$—are ignored in the expansion of f. In the example just given, for instance, the terms of order ω^2 or lower are the only ones considered.

As an example, let us expand the function $f(x) = e^x$ into a Maclaurin series. Since e^x is its own derivative,

$$\frac{d}{dx} e^x = e^x \tag{C.13}$$

and $e^0 = 1$, this generalizes to

$$\left. \frac{d^n}{dx^n} e^x \right|_{x=0} = e^x \big|_{x=0} = e^0 = 1, \tag{C.14}$$

---

[1] R. Feynman. *The Character of Physical Law.* Cambridge, Massachusetts: The M.I.T. Press, 1985, p. 34.

and Eq. (C.2) can be rewritten as follows:

$$\begin{aligned} f(x) = e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \left. \frac{d^n}{dx^n} e^x \right|_{x=0} \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \ldots. \end{aligned} \tag{C.15}$$

Appendix D
Quantum mechanics in a nutshell

D.1 WAVE MECHANICS: THE INTUITIVE APPROACH

According to Felix Bloch's recollections of his student days in Zürich, colloquia were being held alternatively in the departments of the University of Zürich and the Eidgenössische Technische Hochschule:[1]

> Once at the end of a colloquium I heard Debye saying something like: "Schrödinger, you are not working right now on very important problems anyway. Why don't you tell us some time about that thesis of de Broglie, which seems to have attracted some attention?" So in one of the next colloquia,[2] Schrödinger gave a beautifully clear account of how de Broglie associated a wave with a particle and how he could obtain the quantization rules of Niels Bohr and Sommerfeld by demanding that an integer number of waves should be fitted along a stationary orbit. When he had finished, Debye casually remarked that this way of talking was rather childish. As a student of Sommerfeld he had learned that, to deal properly with waves, one had to have a wave equation. It sounded quite trivial and did not seem to make a great impression, but Schrödinger evidently thought a bit more about the idea afterwards. Just a few weeks later,[3] he gave another talk in the colloquium which he started by saying: "My colleague Debye suggested that one should have a wave equation; well I have found one!" And then he told us essentially what he was about to publish under the title "Quantization as an Eigenvalue Problem" as the first paper of a series in the Annalen der Physik.[4]

Schrödinger was an adept of eastern philosophies with their fascination for waves permeating all beings.[5] He considered particles to be merely "epiphenomena" traveling on the crests of the waves. His efforts to explain the behavior of an electron attracted to a proton gave rise to the wave mechanics of the hydrogen atom. Let us reconstruct the mechanics here in an intuitive way. As a starting point, consider a standing sinusoidal wave in one dimension:

$$\Psi = A \sin \frac{2\pi x}{\lambda}, \tag{D.1}$$

---

[1] F. Bloch. "Reminiscences of Heisenberg and the Early Days of Quantum Mechanics." *Physics Today* 29 (1976), p. 23.

[2] According to Moore, this was probably on November 23, 1925. See W. Moore. *Schrödinger, Life and Thought*. Cambridge: Cambridge University Press, 1989.

[3] This must have been in January 1926, after his winter holidays in Arosa.

[4] E. Schrodinger. "Quantization as an Eigenvalue Problem. [Quantisierung als Eigenwertproblem]." *Annalen der Physik* 79(1926), pp. 361–376.

[5] Moore. *Schrödinger, Life and Thought.*

where A is the amplitude, corresponding to the maximal value of the wave function. The constant λ is the wavelength. This measure is equal to the repeat length of the sinusoidal pattern. Whenever the position coordinate x runs through an integer number of λ's, the sine function runs through zero and begins another period. The function Ψ describes a standing wave. To forward this wave in the positive x direction with a constant speed v, we have to substitute x with $x - vt$, yielding the expression for a running sinusoidal wave:

$$\Psi = A \sin\left(\frac{2\pi}{\lambda}(x - vt)\right). \tag{D.2}$$

The equivalent of the wavelength in position space is the period, T, on the time axis. The period measures the time interval (in seconds) for the passage of a complete wavelength. The inverse of the period is the frequency, ν. It is expressed in Hertz ($1Hz = 1s^{-1}$) and counts the number of oscillations per second. These quantities are related as follows:

$$v = \frac{\Delta s}{\Delta t} = \frac{\lambda}{T} = \nu\lambda. \tag{D.3}$$

The wave may thus be rewritten as

$$\Psi = A \sin\left(\frac{2\pi x}{\lambda} - 2\pi\nu t\right). \tag{D.4}$$

Taking the second derivative of the wave function in the x-coordinate yields

$$\frac{d^2}{dx^2}\Psi = -\frac{4\pi^2}{\lambda^2} A \sin\left(\frac{2\pi x}{\lambda} - 2\pi\nu t\right) = -\frac{4\pi^2}{\lambda^2}\Psi. \tag{D.5}$$

The wave thus obeys a second-order differential equation in position space, which is termed the *wave equation*. In three dimensions, this equation reads

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\Psi = -\frac{4\pi^2}{\lambda^2}\Psi. \tag{D.6}$$

Schrödinger compared the motion of an electron near the atomic nucleus to a three-dimensional wave wrapped around the nucleus. To obtain an explicit form of this equation, it was necessary to insert a value for the wavelength. In the beginning of the twentieth century, two apparently unrelated expressions had become available, both involving energies. In 1900, Planck postulated a relationship between the energy and frequency of light, known as *Planck's relation*:

$$E = h\nu. \tag{D.7}$$

On the basis of this expression, Planck could obtain a derivation of Wien's law, which describes the radiation of a black body as a function of temperature. Planck's relation is usually considered to be the starting point of quantum mechanics because it was claimed that energy changes could only be achieved in quanta. In 1905, Einstein applied Planck's relation to the photoelectric effect and came to the conclusion that light has a dual nature: it can be represented as an electromagnetic wave, but also as a stream of particles or photons. On colliding with electrons in a metal, the photons are able to eject electrons from the metal. In that same year, Einstein published his iconic formula expressing the equivalence between mass and energy, where c is the velocity of light:

$$E = mc^2. \tag{D.8}$$

The context of both equations is clearly different, but both are concerned with energies, so let us see what results when we, naively, combine them:

$$h\nu = mc^2. \tag{D.9}$$

Using Eq. (D.3) to replace ν with c/λ, and eliminating c left and right, the right-hand side becomes mc:

$$\frac{hc}{\lambda} = mc^2$$

$$\frac{h}{\lambda} = mc. \tag{D.10}$$

Note that this derivation stands for photons only. It can, however, be generalized to massive particles such as an electron with mass m_e and velocity v, yielding

$$\frac{h}{\lambda} = m_e v = p. \tag{D.11}$$

Here, p is the impulse or momentum. It corresponds to the product of the mass with the velocity and refers to a property of a moving body, as defined in classical Newtonian mechanics. In contrast, the left-hand side contains a wavelength and thus refers to a wavelike phenomenon. This relation is named after Louis de Broglie, who introduced it in his doctoral dissertation in 1924.[6] The connection between waves and particles is a key relation expressing the duality of elementary particles such as electrons. The nature of electrons provides evidence both for wavelike properties, such as interference, and for corpuscular aspects, such as the photoelectric effect. Although duality as a concept seems to defy understanding, the duality relation itself is a simple, practical tool to continue the calculation of the wave equation. The factor preceding Ψ in the right-hand side of Eq. (D.5) may thus be replaced with

$$\frac{4\pi^2}{\lambda^2} = \frac{4\pi^2 p^2}{h^2} = \frac{p^2}{\hbar^2}, \tag{D.12}$$

where we have made use of the symbol $\hbar = h/2\pi$. Newtonian mechanics provides a relationship between momentum and kinetic energy:

$$E_{kin} = \frac{1}{2m} p^2. \tag{D.13}$$

Finally, we can introduce the total energy by taking into account the potential energy, which for the case of the hydrogen atom is based on the Coulomb attraction between the electron and the nucleus:

$$E = E_{kin} + E_{pot} = \frac{1}{2m} p^2 - \frac{e^2}{4\pi\epsilon_0 r}. \tag{D.14}$$

On substitution, we then obtain the famous nonrelativistic wave equation for the hydrogen atom:

$$-\frac{h^2}{8\pi^2 m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\Psi - \frac{e^2}{4\pi\epsilon_0 r}\Psi = E\Psi. \tag{D.15}$$

In this succession of substitutions, the original parameter for the wavelength is finally replaced with the parameter for the total energy, which for the moment is as equally unknown as the wavelength. The so-called mathematical solutions of the equation thus consist of finding the functions Ψ for any value of E. Most of these solutions are not relevant from a physical point of view. What is needed to filter out the physically relevant solutions are additional boundary conditions defining the physical situation of an electron in the attraction of a proton. To derive this condition, we can make use of a statistical

---

[6] L. de Broglie. “Investigations on the Theory of Quanta. [Recherches sur la théorie des Quanta. Thèse de doctorat soutenue à Paris, le 25 Novembre 1924].” *Annales de Physique* III.10 (1925), pp. 22–128.

interpretation to connect the wave back to the particle. In this interpretation, the absolute square $\Psi^*\Psi$ is interpreted as the probability density. In other words, the probability, dP, that the electron is in a volume dV is given by

$$\frac{dP}{dV} = \Psi^*\Psi. \tag{D.16}$$

This relation is known as the *Born relation.* A bound electron should be found in the vicinity of the nucleus, and the corresponding probability should decrease steadily when one is moving away from the nucleus. This implies that the total density, obtained by integrating dP over space, is finite.[7] Functions that obey this criterion are said to be *square integrable.*

$$P = \int dP = \iiint \Psi^*\Psi \, dV < \infty. \tag{D.17}$$

Combining this condition with the wave equation filters out a series of quantized energies that match the energy levels observed in the spectrum of a hydrogen atom. The results rely only on fundamental constants: the mass and charge of the electron, and the proportionality constants in Planck's and Coulomb's law. On a more general level, wave mechanics made it possible to derive from first principles the fundamental atomic units of length and energy, corresponding to the Bohr radius and the Hartree energy, respectively. All these aspects are examined in more detail in Chapter 9 on SO(4) symmetry. Last but not least, the fundamental symmetries of the smallest atom seemed to imply the Aufbau structure of the entire periodic table. This claim is examined in Chapter 13.

The intuitive combination of classical physics (with Newton and Coulomb), statistics (Born), and the particle–wave duality relation (Planck, Einstein, de Broglie) led to the creation of wave mechanics. For further development, a more formal structure had to be established. In the following section, we will examine its basic elements.

D.2 QUANTUM MECHANICS: THE FORMAL STRUCTURE

Several textbooks provide a thorough introduction to quantum mechanics.[8] Here, we limit ourselves to the basic aspects that have a bearing on the main text.

D.2.1 Operators and eigenfunctions

The central concept of quantum mechanics is the *wave function*, which describes the physical state of a system. In fact, all information that can be obtained about the system is included in the wave function. The wave function is the solution of the wave equation, which appears as a linear equation in Ψ, where on the left-hand side the wave function is preceded by a function of coordinates and their derivatives, whereas on the right-hand side it is preceded by the energy constant. Such a type of equation is called an *eigenvalue equation.* The functional factor to the left is the energy operator, or Hamiltonian, $\mathscr{H}$, whereas the energy to the right is the eigenvalue of this operator. The function that solves the equation for a given E is the associated *eigenfunction.* The concise form of the equation thus reads

$$\mathscr{H}\Psi = E\Psi. \tag{D.18}$$

---

[7] In the original paper, Schrödinger stated he was looking for a finite and single-valued function that was twice differentiable. See E. Schrödinger. "Quantization as an Eigenvalue Problem," p. 363.

[8] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Mécanique Quantique.* Paris: Hermann, 1973.

The Hamiltonian is based on the classical energy expression $E = p^2/2m + E_{pot}$. To turn this expression into operator form, the distance dependence in the potential energy part is simply rendered as the radius r, which is given by the coordinate function $\sqrt{x^2 + y^2 + z^2}$. Hence, the operator associated with a coordinate is just the multiplication with that coordinate function:

$$\hat{x} = x; \\ \hat{r} = r. \tag{D.19}$$

Here we use the hat notation to explicitly refer to the operator form. In the kinetic energy part, the momenta are replaced by differential operators in the coordinates:

$$\hat{p}_x = \frac{\hbar}{i} \frac{\partial}{\partial x}. \tag{D.20}$$

The total kinetic energy operator then becomes

$$\frac{1}{2m} \left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 \right) = -\frac{h^2}{8\pi^2 m_e} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right). \tag{D.21}$$

The set of eigenfunctions of the Hamiltonian forms a space known as the *Hilbert space*. Especially, the standard work of Dirac has been instrumental in constructing the algebraic framework of quantum mechanics.[9] For reasons that will become clear in a moment, Dirac denotes the wave function as a *ket* vector in the Hilbert space, or simply a ket, with symbol $|\Psi\rangle$. Using ket vector notation, Schrödinger's wave equation may now be rewritten as

$$\mathscr{H} |\Psi\rangle = E |\Psi\rangle. \tag{D.22}$$

This result can be generalized to any physical observable. In quantum mechanics, a physical observable is represented by an *operator*. The allowed values this observable can acquire are the eigenvalues of the corresponding eigenvalue equation:

$$\hat{A} |a_k\rangle = a_k |a_k\rangle. \tag{D.23}$$

Here, $\hat{A}$ is the operator corresponding to a physical observable and $|a_k\rangle$ is an eigenfunction of this operator with associated eigenvalue a_k. If the system is in the state $|a_k\rangle$, the measurement of the physical property will yield the eigenvalue a_k. But, what if the system is in a different state? Say, for instance, the ground state of the Hamiltonian, $|\Psi_0\rangle$? What is the result of the measurement of the property $\hat{A}$ in this case? To answer this question, we first need to extend the framework to "bra" functions and determine the Hermiticity of quantum mechanical operators.

D.2.2 The bra-ket formalism

The space of eigenvectors has a dual space that Dirac called the *bra* vectors, with symbol $\langle\Psi|$. The bra combines with the ket to form a scalar product or bracket, hence the terminology *bra* and *ket* to denote its constituents. The combination between the bra $\langle\Phi|$ and the ket $|\Psi\rangle$ is written as $\langle\Phi|\Psi\rangle$. This combination is nothing more than the overlap of Φ and Ψ. We may thus write

$$\langle\Phi|\Psi\rangle = \iiint \Phi^* \Psi \, dV. \tag{D.24}$$

Hence, we also have

$$\langle\Phi|\Psi\rangle = \langle\Psi|\Phi\rangle^*. \tag{D.25}$$

---

[9] P. A. M. Dirac. *The Principles of Quantum Mechanics.* Oxford: Clarendon Press, 1930.

The integral of the density $\Psi^*\Psi$ thus corresponds to the bracket of $\langle\Psi|$ and $|\Psi\rangle$. This bracket is the norm of the wave function:

$$|N|^2 = \langle\Psi|\Psi\rangle. \tag{D.26}$$

In accordance with the Born interpretation, this bracket represents the probability that the particle is found in space. By dividing the wave function through N, we obtain a normalized wave function for which the total probability is equal to one. Note that such rescaling of the wave function does not interfere with the eigenvalue equation because this is linear in $|\Psi\rangle$. Henceforth, ket functions are assumed to be normalized.

The duality between bras and kets allows us to determine the action of an operator on the bra functions. Consider the action of an operator $\hat{A}$ on a function $|f\rangle$, which turns this function into some other function $|f'\rangle$:

$$\hat{A}|f\rangle = |f'\rangle. \tag{D.27}$$

Transposing this expression to the bra part would read

$$\langle\hat{A}f| = \langle f'|. \tag{D.28}$$

This expression is rewritten as

$$\langle\hat{A}f| = \langle f|\hat{A}^\dagger. \tag{D.29}$$

Here, the operator $\hat{A}^\dagger$ is called the *adjoint operator* of $\hat{A}$. It is acting on the bras from the right-hand side and is defined as the operator that transforms the bras in a way that is conjugate to the action of the original operator on the kets. Hence, we also have

$$\langle\Phi|\hat{A}^\dagger|\Psi\rangle = \langle\Psi|\hat{A}|\Phi\rangle^*. \tag{D.30}$$

Note that on the right-hand side of this equation, the notation $\langle\Psi|\hat{A}|\Phi\rangle$ is equivalent to $\langle\Psi|\hat{A}\Phi\rangle$. Taking the adjoint of the adjoint returns the original operator, and the adjoint of a product is the transposed product of the adjoints:

$$\begin{aligned}\left(\hat{A}^\dagger\right)^\dagger &= \hat{A};\\ \left(\hat{A}\hat{B}\right)^\dagger &= \hat{B}^\dagger\hat{A}^\dagger.\end{aligned} \tag{D.31}$$

D.2.3 Hermitian and unitary operators

In quantum mechanics we encounter two important special types of operators: *Hermitian* and *unitary*.

Hermitian operators

For Hermitian operators, the adjoint is equal to the operator itself. Hence, the following equations hold:

$$\begin{aligned}\hat{A}^\dagger &= \hat{A};\\ \langle\Phi|\hat{A}|\Psi\rangle &= \langle\Psi|\hat{A}|\Phi\rangle^*.\end{aligned} \tag{D.32}$$

For $\Phi = \Psi$, this equation thus becomes

$$\langle\Psi|\hat{A}|\Psi\rangle = \langle\Psi|\hat{A}|\Psi\rangle^*. \tag{D.33}$$

Now suppose that $|\Psi\rangle$ is an eigenfunction of $\hat{A}$ with eigenvalue a_k. Then we have

$$\langle a_k|\hat{A}|a_k\rangle = a_k\langle a_k|a_k\rangle = a_k;$$
$$\langle a_k|\hat{A}|a_k\rangle^* = a_k^*\langle a_k|a_k\rangle^* = a_k^*. \tag{D.34}$$

In view of Eq. (D.33), it is implied that $a_k = a_k^*$; hence, the eigenvalue of a Hermitian operator is a real number. This is precisely what is required for physical observables, which are determined by quantitative measurements. It can thus be stated:

Operators, corresponding to physical observables, must be Hermitian.

Now consider two eigenstates $|a_k\rangle$ and $|a_l\rangle$ of a Hermitian operator $\hat{A}$, with different eigenvalues:

$$\begin{aligned}\langle a_k|\hat{A}a_l\rangle &= a_l\langle a_k|a_l\rangle \\ &= \langle \hat{A}a_k|a_l\rangle \\ &= a_k\langle a_k|a_l\rangle.\end{aligned} \tag{D.35}$$

Or,

$$(a_k - a_l)\,\langle a_k|a_l\rangle = 0. \tag{D.36}$$

If $a_k \neq a_l$, this result implies that the overlap between the two eigenvectors must vanish. The eigenvectors are said to be orthogonal. A set of eigenvectors that are normalized and also orthogonal to each other form an orthonormal basis:

$$\langle a_k|a_l\rangle = \delta_{kl}. \tag{D.37}$$

The action of an operator in a basis is described by the brackets $\langle a_k|\hat{O}a_l\rangle$. These brackets are called the *matrix elements* O_{kl} of a matrix $\mathbb{O}$, describing the action of the operator. For $\hat{O}$ being a Hermitian operator, the associated matrix is a Hermitian matrix:

$$O_{ij} = O_{ji}^*. \tag{D.38}$$

Unitary operators

Another important class of operators includes the unitary operators, which we denote as $\hat{U}$. They fulfill the requirement that the adjoint is equal to the inverse:

$$\hat{U}^\dagger = \hat{U}^{-1};$$
$$\langle \hat{U}\Phi|\hat{U}\Psi\rangle = \langle \Phi|\hat{U}^{-1}\hat{U}\Psi\rangle = \langle \Phi|\Psi\rangle. \tag{D.39}$$

Hence, unitary operators conserve an orthonormal basis. The matrix elements of unitary operators thus obey:

$$\langle \Phi|\hat{U}^{-1}|\Psi\rangle = \langle \Psi|\hat{U}|\Phi\rangle^*. \tag{D.40}$$

This implies that unitary operators are represented by unitary matrices, defined in Appendix B. As we demonstrate in Chapter 6, symmetry transformations are unitary operators. Now suppose that $|\Psi_u\rangle$ is an eigenfunction of a unitary operator, with eigenvalue u:

$$\hat{U}|\Psi_u\rangle = u|\Psi_u\rangle;$$
$$\langle \Psi_u|\hat{U}^\dagger = u^*\langle \Psi_u|. \tag{D.41}$$

Then, we have

$$\langle \Psi_u | \hat{U}^\dagger \hat{U} | \Psi_u \rangle = \langle \Psi_u | \Psi_u \rangle = u^* u = 1. \tag{D.42}$$

This result implies that eigenvalues of unitary operators must be unimodular. Finally, it can be proven that if $\hat{A}$ is Hermitian, then $\hat{T} = \exp(i\hat{A})$ is unitary. We have

$$\hat{T}^\dagger = e^{-i\hat{A}^\dagger} = e^{-i\hat{A}}. \tag{D.43}$$

This implies: $\hat{T}\hat{T}^\dagger = \hat{T}^\dagger \hat{T} = \hat{E}$, where $\hat{E}$ is the unit operator. This proves that the adjoint is the inverse operator—meaning, $\hat{T}$ is unitary:

$$\hat{T}^\dagger = \hat{T}^{-1}. \tag{D.44}$$

D.2.4 The spectral decomposition postulate

We can now return to our question: What is the result of a measurement of property $\hat{A}$ on a state described by an eigenfunction of another operator—say, $|\Psi_0\rangle$ of $\mathscr{H}$? We first expand this eigenstate in the basis of the $\hat{A}$ operator, with c_k being the expansion coefficient:

$$|\Psi_0\rangle = \sum_k c_k |a_k\rangle. \tag{D.45}$$

The expansion coefficient can also be expressed as a bracket because it corresponds to the overlap between $|\Psi_0\rangle$ with the basis function $|a_k\rangle$:

$$c_k = \langle a_k | \Psi_0 \rangle. \tag{D.46}$$

In view of the conjugation relation between bra and ket, the corresponding bra is then given by

$$\langle \Psi_0 | = \sum_k \langle a_k | c_k^*. \tag{D.47}$$

Since the basis is orthonormal and $|\Psi_0\rangle$ is normalized, we have

$$\langle \Psi_0 | \Psi_0 \rangle = \sum_{kl} c_l^* c_k \langle a_l | a_k \rangle = \sum_{kl} c_l^* c_k \delta_{lk} = \sum_k |c_k|^2 = 1. \tag{D.48}$$

As indicated earlier, the measurement can yield only one of the eigenvalues of the corresponding operator. Now, according to the postulate of *spectral decomposition*, the probability of measuring the particular eigenvalue a_k, $P(a_k)$, is given by

$$P(a_k) = |c_k|^2. \tag{D.49}$$

Hence, if the measurement is applied to a large set of systems, the outcome is a spectrum of discrete eigenvalues of the operator, which is distributed according to the statistics in Eq. (D.49). Conversely, the average value of the measurement can be evaluated from the bracket $\langle \Psi_0 | \hat{A} | \Psi_0 \rangle$ as

$$\begin{aligned}
\langle \hat{A} \rangle &= \langle \Psi_0 | \hat{A} | \Psi_0 \rangle \\
&= \left\langle \sum_k c_k a_k \middle| \sum_l \left(\hat{A}\, c_l a_l \right) \right\rangle \\
&= \sum_{k,l} c_k^* c_l a_l \langle a_k | a_l \rangle \\
&= \sum_{k,l} c_k^* c_l a_l \delta_{k,l} \\
&= \sum_k a_k |c_k|^2.
\end{aligned} \tag{D.50}$$

In some cases, special results may be obtained depending on the operators involved, as we will see in the next section.

D.3 OPERATOR COMMUTATIONS AND UNCERTAINTY RELATIONS

From the point of view of classical physics, the process of measurement in quantum mechanics is perhaps the most problematic and least comprehensible aspect of the new physics. The spectral decomposition theorem tells us that the outcome of the measurement for a physical property associated with an operator $\hat{A}$ on a quantum state of an operator $\mathscr{H}$ is unpredictable and can be known only with a certain probability. Moreover, when the measurement has been performed, the system has adopted the corresponding eigenstate of $\hat{\mathrm{A}}$. This implies that, immediately afterward, remeasuring the same property again yields the same result.

There is, however, a possibility that two physical quantities are compatible (i.e., that a system can be in a state that is simultaneously an eigenstate of both). The compatibility requirement is simply that both associated operators commute. Consider as an example the commutator of $\mathscr{H}$ and $\hat{\mathrm{A}}$. It is indicated by square brackets and is defined as follows:

$$\left[\hat{\mathrm{A}}, \mathscr{H}\right] = \hat{\mathrm{A}}\mathscr{H} - \mathscr{H}\hat{\mathrm{A}}. \tag{D.51}$$

Note that the product of operators simply means that they should be acting on the ket function to the right in consecutive order. Now let us consider the specific case when this commutator vanishes: $\left[\hat{A}, \mathscr{H}\right] = 0$. For simplicity, let us further assume that the eigenvalues of $\hat{A}$ are non-degenerate. In this case, we can write

$$\begin{aligned} \mathscr{H}\hat{\mathrm{A}}|a_k\rangle &= a_k \mathscr{H}|a_k\rangle \\ &= \hat{\mathrm{A}}\mathscr{H}|a_k\rangle. \end{aligned} \tag{D.52}$$

Hence,

$$\hat{\mathrm{A}}\mathscr{H}|a_k\rangle = a_k \mathscr{H}|a_k\rangle. \tag{D.53}$$

Eq. (D.53) shows that $\mathscr{H}|a_k\rangle$ is an eigenfunction of $\hat{A}$ with the same eigenvalue as $|a_k\rangle$. But, because there is only one such eigenfunction in case of nondegeneracy, we must have

$$\mathscr{H}|a_k\rangle \sim |a_k\rangle. \tag{D.54}$$

This proves $|a_k\rangle$ is also an eigenfunction of $\mathscr{H}$.

If the commutator is nonzero, the operators are no longer compatible. In this case, a given state cannot, in general, be an eigenstate of both. An important case of non-commutation is between the momentum and the coordinate:

$$\left[\hat{p}_x, \hat{x}\right] = \frac{\hbar}{i}\left(\frac{\partial}{\partial x}x - x\frac{\partial}{\partial x}\right) = \frac{\hbar}{i}. \tag{D.55}$$

Hence, both the position and the momentum of a particle cannot have a definite eigenvalue simultaneously. Let us examine both aspects for the extreme cases of a free particle moving along the x direction versus an immobile particle at a given position. The eigenvalue equation for the momentum operator is given by

$$\hat{p}_x|\Psi\rangle = p_0|\Psi\rangle. \tag{D.56}$$

The solution of this equation reads

$$|\Psi\rangle = e^{ip_0x/\hbar} = e^{ik_0x}. \tag{D.57}$$

Here, $k_0 = p_0/\hbar$. The resulting eigenfunction with definite momentum p_0 is a continuous wave along the x direction, that has everywhere the same density.[10] Hence, the position

---

[10] Since this wave has the same amplitude everywhere in space, it cannot be normalized.

of the particle is completely undetermined. Conversely, the eigenvalue equation for the position operator is given by

$$\hat{x}|\Phi\rangle = x_0|\Phi\rangle. \tag{D.58}$$

The resulting eigenfunction is a function that exists only at the point $x = x_0$, and it vanishes everywhere else. This function can be defined as the limiting case of a Gaussian function that is becoming infinitely sharp:

$$\delta(x - x_0) = \lim_{a \to 0} \frac{1}{a\sqrt{\pi}}\, e^{-(x-x_0)^2/a^2}. \tag{D.59}$$

Although in this way the position of the particle is determined accurately at $x = x_0$, its momentum is completely undetermined. Indeed, if we decompose this function in plane waves by means of a Fourier transformation, the entire spectral range for p is required. This type of inverse relationship between coordinates and momenta has found an accurate expression in the Heisenberg uncertainty principle. One of the manifestations of this principle states that the product of the uncertainty of the position of a particle, Δx, multiplied by the uncertainty of its momentum, Δp_x, is bounded from below by Planck's constant according to

$$\Delta x.\Delta p_x \leq \frac{\hbar}{2}. \tag{D.60}$$

Note that this relationship is not an additional external constraint, but is an intrinsic property resulting from the noncommutation of the corresponding operators. It is noteworthy that Lie algebras as well, which are a common thread in this book, are based essentially on commutation relations.

D.4 TIME DEPENDENCE

So far, we have not been concerned with the time dependence of the wave function. Instead, we have only looked at stationary states of systems subject to time-independent interactions. Returning to the classical expression for a propagating wave in Eq. (D.4), differentiation with respect to time yields the following second-order equation:

$$\frac{\partial^2}{\partial t^2}\Psi = -4\pi^2\nu^2\Psi. \tag{D.61}$$

The transition to a quantum wave equation in this case is not straightforward because the frequency, and hence the energy, appears as a second power. However, we can turn this equation into a first-order differential equation, the solutions of which are the complex combinations of the cosine and sine waves:

$$i\frac{\partial}{\partial t}\Psi = 2\pi\nu\Psi, \tag{D.62}$$

with

$$\Psi = e^{i(2\pi x/\lambda - 2\pi\nu t)}. \tag{D.63}$$

Making use of Planck's equation, we now obtain

$$i\hbar\frac{\partial}{\partial t}\Psi = E\Psi. \tag{D.64}$$

The quantum mechanical equation, which is known as Schrödinger's time-dependent equation, is then obtained by replacing the energy with the corresponding Hamiltonian operator:

$$i\hbar\frac{\partial}{\partial t}\Psi = \mathscr{H}\Psi. \tag{D.65}$$

This equation predicts the evolution of the state of the system under time-dependent interactions that may occur in the Hamiltonian.[11] In case the Hamiltonian is time-independent, Schrödinger's first wave equation applies and the right-hand side of Eq. (D.65) can be replaced again wither $E\Psi$, where this time E is the stationary eigenenergy for which Ψ is an eigenfunction. Integration of the time-dependent equation then yields

$$\Psi(t) = \Psi(t_0)e^{-iE(t-t_0)/\hbar}. \tag{D.66}$$

Hence, time dependence in this case is limited to a dynamic phase factor. The phase is pulsating with a frequency that corresponds to E/h, exactly as in Planck's law.

---

[11] Note that the momentum operator is preceded by $\hbar/i$ whereas the time operator is preceded by $i\hbar$. These factors differ by a minus sign, which can be traced back to the minus sign in the argument $x - vt$, which appears in the description of the running wave.

Appendix E
Determinant of an orthogonal matrix

The derivation in Eq. (5.29) is a corollary of a more general relationship for the derivative of a determinant known as *Jacobi's formula.* The *minor* associated with the element a_{ij} of a square matrix $\mathbb{A}$ is defined as the determinant of the submatrix, which is obtained by removing row i and column j from $\mathbb{A}$. The minor multiplied by the phase factor $(-1)^{i+j}$ is called the *cofactor* and is denoted as α_{ij}. As an example, for a 3 × 3 matrix, we have

$$\begin{aligned} \alpha_{11} &= +\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}; \\ \alpha_{12} &= -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -a_{21}a_{33} + a_{23}a_{31}; \\ \alpha_{13} &= +\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31}. \end{aligned} \tag{E.1}$$

According to the well-known Laplace expansion, the determinant of the full matrix can then be obtained by the scalar product of a given row with its respective cofactors:

$$\det \mathbb{A} = a_{11}\alpha_{11} + a_{12}\alpha_{12} + a_{13}\alpha_{13}. \tag{E.2}$$

A more general result reads

$$\sum_k a_{ik}\alpha_{jk} = \sum_k \alpha_{ik}a_{jk} = \delta_{ij} \det \mathbb{A}. \tag{E.3}$$

We can now define the classical adjoint or *adjugate* matrix $\mathbb{B} = \mathrm{adj}\,(\mathbb{A})$ as the transpose of the matrix of cofactors:

$$b_{ij} = \alpha_{ji}. \tag{E.4}$$

The result in Eq. (E.3) can thus be rewritten as

$$\mathbb{A} \times \mathbb{B} = \mathbb{B} \times \mathbb{A} = (\det \mathbb{A})\,\mathbb{I}. \tag{E.5}$$

The Laplace result $\det \mathbb{A} = \sum_k a_{ik}\alpha_{ik}$ further shows that the determinant is a linear function of any given matrix element, because the only term in the determinant containing a given a_{il} is the $k = l$ term in the Laplace expansion, with the expansion coefficient being the corresponding cofactor α_{il}, which is a polynomial. Hence, we can write

$$\frac{\partial \det \mathbb{A}}{\partial a_{ij}} = \alpha_{ij}. \tag{E.6}$$

Or, for the total derivative, we have

$$\begin{aligned} d\det\mathbb{A} &= \sum_{i,j} \frac{\partial \det\mathbb{A}}{\partial a_{ij}}\, da_{ij} \\ &= \sum_{i,j} \alpha_{ij}\, da_{ij} \\ &= \sum_{ij} b_{ji}\, da_{ij} \\ &= \sum_{j} (\mathbb{B}\times d\mathbb{A})_{jj} \\ &= \mathrm{Tr}\,(\mathbb{B}\times d\mathbb{A})\,. \end{aligned} \tag{E.7}$$

Or, in summary,

$$d\det\mathbb{A} = \mathrm{Tr}\left(\mathrm{adj}\,(\mathbb{A})\times d\mathbb{A}\right). \tag{E.8}$$

This is the Jacobi formula. Now we can derive the corollary we used in Chapter 5. If the matrix under consideration is nonsingular, Eq. (E.5) implies a simple relationship between the inverse matrix and the adjugate:

$$\mathbb{A}^{-1} = \frac{\mathrm{adj}\,(\mathbb{A})}{\det\mathbb{A}}. \tag{E.9}$$

Inserting this in Jacobi's formula we obtain

$$d\det\mathbb{A} = \det\mathbb{A}\;\mathrm{Tr}\left(\mathbb{A}^{-1}\times d\mathbb{A}\right). \tag{E.10}$$

At this point, we can express matrix $\mathbb{A}$ in exponential form, in line with Eq. (5.22):

$$\begin{aligned} \mathbb{A} &= \exp\mathbb{C}; \\ \mathbb{A}^{-1} &= \exp\,(-\mathbb{C}). \end{aligned} \tag{E.11}$$

Here, $\mathbb{C}$ can be any complex square matrix of the same dimension as $\mathbb{A}$. The derivation is thus not limited to skew-symmetric matrices, as it is in case of orthogonal matrices. This also implies

$$d\mathbb{A} = \exp\mathbb{C}\; d\mathbb{C}. \tag{E.12}$$

Combining these results together, and using $\mathrm{Tr}\,(d\mathbb{C}) = d\,\mathrm{Tr}\,\mathbb{C}$, we can rewrite Eq. (E.10) as

$$\begin{aligned} d\det\mathbb{A} &= \det\mathbb{A}\;\mathrm{Tr}\left(\exp\,(-\mathbb{C}+\mathbb{C})\times d\mathbb{C}\right) \\ &= \det\mathbb{A}\;\mathrm{Tr}\,(\mathbb{I}\times d\mathbb{C}) \\ &= \det\mathbb{A}\; d\,\mathrm{Tr}\,\mathbb{C} \end{aligned} \tag{E.13}$$

or

$$d\ln(\det\mathbb{A}) = d\,\mathrm{Tr}\,\mathbb{C}. \tag{E.14}$$

Integration of this formula finally yields the desired result:

$$\det\mathbb{A} = \exp\,(\mathrm{Tr}\,\mathbb{C}), \tag{E.15}$$

where $\mathbb{A} = \exp\mathbb{C}$. The final result thus reads

$$\det\left(\exp\mathbb{C}\right) = \exp\,(\mathrm{Tr}\,\mathbb{C}). \tag{E.16}$$

Appendix F
Lie bracket

The *Lie bracket* was introduced in Definition 5.3 as the combination rule of a Lie algebra. But, where does this definition come from? The full answer to this question takes us to a special branch of mathematics called *differential geometry*. Several books have developed this subject with endless care for generality and abstraction. A lively account can be found in the book by Michael Spivak.[1] In this appendix, however, let us use a more intuitive approach aimed at maximal understanding with minimal introduction of new concepts. Our discussion will be focused on the particular case of the SO(3) group, but the extension to unitary groups is straightforward. The advantage of working with the SO(3) group is that we can explain the concept of the Lie bracket by drawing vectors on the surface of a globe (Figure F.1).

A Lie operator is a derivative of a symmetry operation, and as such it has a vectorial meaning. It may be represented as a tangent vector to the surface of a sphere in the direction of the rotation. As an example, an infinitesimal rotation around the z-axis, as described by the $\hat{X}_3$ operator, can be represented in every point of the globe by a small displacement vector, tangent to the parallel circles on the globe and oriented in the direction of increasing longitude. In Figure F.1, such a vector is drawn at a point somewhere in the Atlantic Ocean on the Equator, midway between Brazil and Africa. Note, however, that such vectors are defined at all points of the sphere with the exception of the rotational poles. They are said to form a *vector field.* At each point of the globe, the vector field tells you where to go to and at what speed.

The action of a second Lie operator in this vector field then means *to take the derivative of the vector field corresponding to the first operator*. As an example, to act on our equatorial vector by an infinitesimal rotation about the x-axis, represented by $\hat{X}_1$, means to rotate this vector in the direction of the Tropic of Cancer, along the dashed circle. As a result, the rotated vector (shown in dark gray) makes a small angle with the original equatorial direction (shown in light gray), and this gives rise to a vectorial difference, which for an infinitesimal rotation angle corresponds to the differential of the original vector. The derivative of a vector field thus evaluates the change of one vector field along the flow of another vector field.

Let us represent this right-justified combination of two operators by the symbol $\left[\hat{X}_1, \hat{X}_3\right]$. As indicated, the meaning of this combination rule is, precisely, to differentiate the $\hat{X}_3$ operator by an infinitesimal rotation about the x direction:

$$\left[\hat{X}_1, \hat{X}_3\right] = \lim_{d\omega \to 0} \frac{\hat{R}_x(d\omega) - \hat{E}}{d\omega} \hat{X}_3. \tag{F.1}$$

---

[1] M. Spivak. *A Comprehensive Introduction to Differential Geometry*, vol. I, 3rd ed. Houston, TX: Publish or Perish 1999.

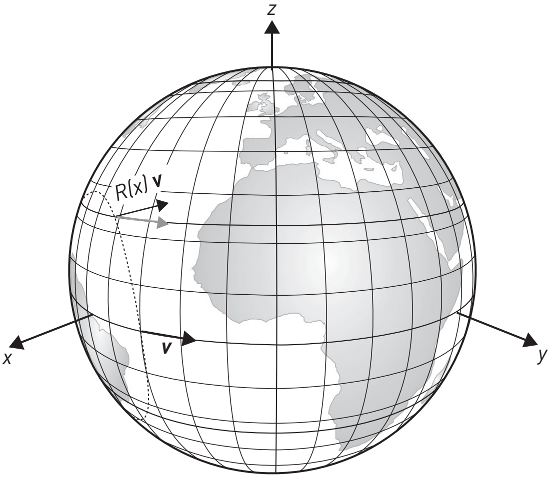


FIGURE F.1 The vector **v** represents the displacement of a point on the Equator by a rotation around the z-axis. $R(x)\,\mathbf{v}$ indicates the action of a second rotation around the x-axis on **v**.

To work out this expression, we must determine how to take the derivative of a Lie operator. As we derived in Chapter 4, the operators can be expressed in matrix form by multiplying the $\mathbb{X}$ matrices to the left by a row vector of the coordinate functions, and to the right by the column vector of their derivative operators:

$$\hat{X}_i = \begin{bmatrix} x & y & z \end{bmatrix} \mathbb{X}_i \begin{bmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial z} \end{bmatrix}. \tag{F.2}$$

We have not yet determined how the differential operators in the column vector on the right transform under rotation. For simplicity, let us study the functional transformations of the differential operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ under rotation around the positive z-axis.[2] To find the transformed operators, we have to work out expressions such as $\frac{\partial}{\partial x'}$, where x' is the rotated form of the x-coordinate function. The chain rule provides a convenient way to relate both coordinate functions:

$$\frac{\partial}{\partial x'} = \frac{\partial x}{\partial x'}\frac{\partial}{\partial x} + \frac{\partial y}{\partial x'}\frac{\partial}{\partial y}. \tag{F.3}$$

From Eqs. (3.36) and (3.37),we already have the primed functions as linear combinations of the unprimed ones:

$$x' = x\cos\alpha + y\sin\alpha; \tag{F.4}$$

$$y' = -x\sin\alpha + y\cos\alpha. \tag{F.5}$$

By inverting these equations, we obtain

$$x = x'\cos\alpha - y'\sin\alpha; \tag{F.6}$$

$$y = x'\sin\alpha + y'\cos\alpha. \tag{F.7}$$

The partial derivatives, needed in the chain rule, can now be obtained by direct derivation:

$$\frac{\partial x}{\partial x'} = \quad \cos\alpha; \tag{F.8}$$

---

[2] This derivation follows the treatment by Ceulemans. See A. Ceulemans. *Group Theory Applied to Chemistry*. Dordrecht: Springer, 2013, p. 8–9.

$$\frac{\partial y}{\partial x'} = \quad \sin\alpha; \tag{F.9}$$

$$\frac{\partial x}{\partial y'} = -\sin\alpha; \tag{F.10}$$

$$\frac{\partial y}{\partial y'} = \quad \sin\alpha. \tag{F.11}$$

Hence, the transformation of the derivatives is entirely similar to the transformation of the x- and y-functions themselves:

$$\hat{R}\begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix}\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}. \tag{F.12}$$

This result is easily generalized to a general element of the SO(3) group:

$$\hat{R}\begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial y} \end{bmatrix}\mathbb{R}\,(\omega\mathbf{n}). \tag{F.13}$$

Alternatively, we can swap this matrix transformation to a column format, which is what we need:

$$\hat{R}\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} = \mathbb{R}^{\mathrm{T}}\,(\omega\mathbf{n})\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix}. \tag{F.14}$$

The result of a general symmetry operation on the operators then follows immediately from the rotation of the row and column vectors, with the $\mathbb{X}$ matrix in the middle being a constant:

$$\hat{R}\,\hat{X}_i = \begin{bmatrix} x & y & z \end{bmatrix}\mathbb{R}\,(\omega\mathbf{n})\;\mathbb{X}_i\,\mathbb{R}^{\mathrm{T}}\,(\omega\mathbf{n})\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix}. \tag{F.15}$$

Remember that the infinitesimal generator of a general rotation is a linear combination of the three generators of the SO(3) group, with the directional cosines as coefficients. The resulting matrix is precisely the $\mathbb{Z}$ matrix:

$$\left.\frac{d}{d\omega}\mathbb{R}\,(d\omega\mathbf{n})\right|_{\omega=0} = n_x\mathbb{X}_1 + n_y\mathbb{X}_2 + n_y\mathbb{X}_3 = \mathbb{Z}. \tag{F.16}$$

Let us represent the corresponding operator form as $\hat{Z}$. As shown in the vectorial representation, the action of this operator on $\hat{X}_i$ is to take the derivative of the $\hat{X}_i$ vector field by rotating it through an infinitesimal angle $d\omega$ about an axis oriented along the $\mathbf{n}$ direction:

$$\begin{aligned} \left[\hat{Z},\hat{X}_i\right] &= \lim_{d\omega\to 0}\frac{\hat{R}\,(d\omega\mathbf{n}) - \hat{E}}{d\omega}\hat{X}_i \\ &= \lim_{d\omega\to 0}\frac{1}{d\omega}\begin{bmatrix} x & y & z \end{bmatrix}\left\{\mathbb{R}\,(d\omega\mathbf{n})\,\mathbb{X}_i\,\mathbb{R}^{\mathrm{T}}\,(d\omega\mathbf{n}) - \mathbb{X}_i\right\}\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix}. \end{aligned} \tag{F.17}$$

The infinitesimal rotation takes us to the immediate vicinity of the unit element and is given by

$$\mathbb{R}\,(d\omega\mathbf{n}) = \mathbb{I} + d\omega\mathbb{Z}; \tag{F.18}$$

$$\mathbb{R}^{\mathrm{T}}(d\omega\mathbf{n}) = \mathbb{I} + d\omega\mathbb{Z}^{\mathrm{T}} = \mathbb{I} - d\omega\mathbb{Z}. \tag{F.19}$$

Here we have used the fact that $\mathbb{Z}$ is an antisymmetric matrix. Combination of these results in the matrix product yields

$$\mathbb{R}(d\omega\mathbf{n})\,\mathbb{X}_i\,\mathbb{R}^{\mathrm{T}}(d\omega\mathbf{n}) - \mathbb{X}_i = d\omega\mathbb{Z}\mathbb{X}_i - d\omega\mathbb{X}_i\mathbb{Z} + \mathcal{O}(2). \tag{F.20}$$

Here, $\mathcal{O}(2)$ is the second-order contribution, which may be neglected. Introducing this result in the product rule leads to

$$\left[\hat{Z},\hat{X}_i\right] = \begin{bmatrix} x & y & z \end{bmatrix}\{\mathbb{Z}\mathbb{X}_i - \mathbb{X}_i\mathbb{Z}\}\begin{bmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial z} \end{bmatrix}. \tag{F.21}$$

To transform this matrix expression back into an operator expression, we use the result from Chapter 4, §4.3.2, that the product of the column of coordinate derivatives multiplied by the row of coordinate functions be equal to the unit matrix:

$$\begin{bmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial z} \end{bmatrix}\begin{bmatrix} x & y & z \end{bmatrix} = \mathbb{I}. \tag{F.22}$$

Inserting this result between the $\mathbb{Z}$ and $\mathbb{X}_i$ matrices allows us to change the matrix forms back to operator forms, finally yielding

$$\left[\hat{Z},\hat{X}_i\right] = \hat{Z}\hat{X}_i - \hat{X}_i\hat{Z}. \tag{F.23}$$

The product of Lie operators thus naturally appears as a commutator, which is the famous Lie bracket.

Appendix G

Laplacian in radial and angular momentum

A tedious but straightforward procedure to find the radial and angular parts of the Laplacian boils down to expressing $\hat{L}^2$ in its Cartesian components using Eq. (5.66) and working out the brackets:

$$\begin{aligned}\hat{L}^2 &= \left|\hat{\mathbf{L}}\right|^2 = \hat{L}_x^2+\hat{L}_y^2+\hat{L}_z^2\\ &= -\hbar^2\left[\underbrace{\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)^2}_{A}+\underbrace{\left(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z}\right)^2}_{B}+\underbrace{\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)^2}_{C}\right].\end{aligned}\tag{G.1}$$

The first term, denoted A, equals

$$\begin{aligned}A &= \left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)\\ &= \underbrace{\left(y\frac{\partial}{\partial z}\right)\left(y\frac{\partial}{\partial z}\right)}_{A1}-\underbrace{\left(y\frac{\partial}{\partial z}\right)\left(z\frac{\partial}{\partial y}\right)}_{A2}-\underbrace{\left(z\frac{\partial}{\partial y}\right)\left(y\frac{\partial}{\partial z}\right)}_{A3}+\underbrace{\left(z\frac{\partial}{\partial y}\right)\left(z\frac{\partial}{\partial y}\right)}_{A4},\end{aligned}\tag{G.2}$$

with

$$\begin{aligned}A1 &= y\frac{\partial y}{\partial z}\frac{\partial}{\partial z}+y^2\frac{\partial}{\partial z}\frac{\partial}{\partial z} = 0+y^2\frac{\partial^2}{\partial z^2};\\ A2 &= y\frac{\partial z}{\partial z}\frac{\partial}{\partial y}+yz\frac{\partial}{\partial z}\frac{\partial}{\partial y} = y\frac{\partial}{\partial y}+yz\frac{\partial^2}{\partial z\partial y};\\ A3 &= z\frac{\partial y}{\partial y}\frac{\partial}{\partial z}+zy\frac{\partial}{\partial y}\frac{\partial}{\partial z} = z\frac{\partial}{\partial z}+zy\frac{\partial^2}{\partial y\partial z};\\ A4 &= z\frac{\partial z}{\partial y}\frac{\partial}{\partial y}+z^2\frac{\partial}{\partial y}\frac{\partial}{\partial y} = 0+z^2\frac{\partial^2}{\partial y^2}.\end{aligned}\tag{G.3}$$

Substituting Eq. (G.3) in Eq. (G.2) yields

$$A = y^2\frac{\partial^2}{\partial z^2}-y\frac{\partial}{\partial y}-yz\frac{\partial^2}{\partial z\partial y}-z\frac{\partial}{\partial z}-zy\frac{\partial^2}{\partial y\partial z}+z^2\frac{\partial^2}{\partial y^2}.\tag{G.4}$$

The second and third term in Eq. (G.1), denoted B and C, are found by twice performing the cyclic permutation of the indices $x \to y \to z \to x$ in Eq. G.4 yielding

$$B = z^2\frac{\partial^2}{\partial x^2}-z\frac{\partial}{\partial z}-zx\frac{\partial^2}{\partial x\partial z}-x\frac{\partial}{\partial x}-xz\frac{\partial^2}{\partial z\partial x}+x^2\frac{\partial^2}{\partial z^2};\tag{G.5}$$

$$C = x^2\frac{\partial^2}{\partial y^2}-x\frac{\partial}{\partial x}-xy\frac{\partial^2}{\partial y\partial x}-y\frac{\partial}{\partial y}-yx\frac{\partial^2}{\partial x\partial y}+y^2\frac{\partial^2}{\partial x^2}.\tag{G.6}$$

Substituting these expressions for A, B, and C in Eq. (G.1) gives

$$\begin{aligned}\hat{L}^2 = -\hbar^2 \Bigg[& y^2\frac{\partial^2}{\partial z^2} - y\frac{\partial}{\partial y} - yz\frac{\partial^2}{\partial z\partial y} - z\frac{\partial}{\partial z} - zy\frac{\partial^2}{\partial y\partial z} + z^2\frac{\partial^2}{\partial y^2} + z^2\frac{\partial^2}{\partial x^2} - z\frac{\partial}{\partial z} - zx\frac{\partial^2}{\partial x\partial z} \\ & - x\frac{\partial}{\partial x} - xz\frac{\partial^2}{\partial z\partial x} + x^2\frac{\partial^2}{\partial z^2} + x^2\frac{\partial^2}{\partial y^2} - x\frac{\partial}{\partial x} - xy\frac{\partial^2}{\partial y\partial x} - y\frac{\partial}{\partial y} - yx\frac{\partial^2}{\partial x\partial y} + y^2\frac{\partial^2}{\partial x^2}\Bigg].\end{aligned} \tag{G.7}$$

Now, let us compare this result with the operator form of $r^2p^2 - (\mathbf{r}\cdot\mathbf{p})^2$ from Eq. (5.134). Given that $\hat{r}^2 = x^2 + y^2 + z^2$ and $\hat{p}^2 = -\hbar^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$, we have

$$\begin{aligned}\hat{r}^2\hat{p}^2 &= -\hbar^2\left(x^2+y^2+z^2\right)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \\ &= -\hbar^2\Bigg[x^2\frac{\partial^2}{\partial x^2} + x^2\frac{\partial^2}{\partial y^2} + x^2\frac{\partial^2}{\partial z^2} + y^2\frac{\partial^2}{\partial x^2} + y^2\frac{\partial^2}{\partial y^2} + y^2\frac{\partial^2}{\partial z^2} \\ &\qquad + z^2\frac{\partial^2}{\partial x^2} + z^2\frac{\partial^2}{\partial y^2} + z^2\frac{\partial^2}{\partial z^2}\Bigg].\end{aligned} \tag{G.8}$$

Also, $\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} = \hbar/i\left(x\partial/\partial x + y\partial/\partial y + z\partial/\partial z\right)$, and thus

$$\begin{aligned}\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 &= -\hbar^2\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right)\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right) \\ &= -\hbar^2\Bigg[\left(x\frac{\partial}{\partial x}x\frac{\partial}{\partial x}\right) + \left(x\frac{\partial}{\partial x}y\frac{\partial}{\partial y}\right) + \left(x\frac{\partial}{\partial x}z\frac{\partial}{\partial z}\right) + \left(y\frac{\partial}{\partial y}x\frac{\partial}{\partial x}\right) \\ &\qquad + \left(y\frac{\partial}{\partial y}y\frac{\partial}{\partial y}\right) + \left(y\frac{\partial}{\partial y}z\frac{\partial}{\partial z}\right) + \left(z\frac{\partial}{\partial z}x\frac{\partial}{\partial x}\right) + \left(z\frac{\partial}{\partial z}y\frac{\partial}{\partial y}\right) + \left(z\frac{\partial}{\partial z}z\frac{\partial}{\partial z}\right)\Bigg] \\ &= -\hbar^2\Bigg[\left(x\frac{\partial x}{\partial x}\frac{\partial}{\partial x} + x^2\frac{\partial}{\partial x}\frac{\partial}{\partial x}\right) + \left(x\frac{\partial y}{\partial x}\frac{\partial}{\partial y} + xy\frac{\partial}{\partial x}\frac{\partial}{\partial y}\right) + \left(x\frac{\partial z}{\partial x}\frac{\partial}{\partial z} + xz\frac{\partial}{\partial x}\frac{\partial}{\partial z}\right) \\ &\qquad + \left(y\frac{\partial x}{\partial y}\frac{\partial}{\partial x} + yx\frac{\partial}{\partial y}\frac{\partial}{\partial x}\right) + \left(y\frac{\partial y}{\partial y}\frac{\partial}{\partial y} + y^2\frac{\partial}{\partial y}\frac{\partial}{\partial y}\right) + \left(y\frac{\partial z}{\partial y}\frac{\partial}{\partial z} + yz\frac{\partial}{\partial y}\frac{\partial}{\partial z}\right) \\ &\qquad + \left(z\frac{\partial x}{\partial z}\frac{\partial}{\partial x} + zx\frac{\partial}{\partial z}\frac{\partial}{\partial x}\right) + \left(z\frac{\partial y}{\partial z}\frac{\partial}{\partial y} + zy\frac{\partial}{\partial z}\frac{\partial}{\partial y}\right) + \left(z\frac{\partial z}{\partial z}\frac{\partial}{\partial z} + z^2\frac{\partial}{\partial z}\frac{\partial}{\partial z}\right)\Bigg] \\ &= -\hbar^2\Bigg[x\frac{\partial}{\partial x} + x^2\frac{\partial^2}{\partial x^2} + xy\frac{\partial^2}{\partial x\partial y} + xz\frac{\partial^2}{\partial x\partial z} + yx\frac{\partial^2}{\partial y\partial x} + y\frac{\partial}{\partial y} \\ &\qquad + y^2\frac{\partial^2}{\partial y^2} + yz\frac{\partial^2}{\partial y\partial z} + zx\frac{\partial^2}{\partial z\partial x} + zy\frac{\partial^2}{\partial z\partial y} + z\frac{\partial}{\partial z} + z^2\frac{\partial^2}{\partial z^2}\Bigg].\end{aligned} \tag{G.9}$$

Combining the last two equations yields

$$\begin{aligned}\hat{r}^2\hat{p}^2 - \left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 = -\hbar^2\Bigg[& x^2\frac{\partial^2}{\partial y^2} + x^2\frac{\partial^2}{\partial z^2} + y^2\frac{\partial^2}{\partial x^2} + y^2\frac{\partial^2}{\partial z^2} + z^2\frac{\partial^2}{\partial x^2} + z^2\frac{\partial^2}{\partial y^2} - x\frac{\partial}{\partial x} \\ & - xy\frac{\partial^2}{\partial x\partial y} - xz\frac{\partial^2}{\partial x\partial z} - yx\frac{\partial^2}{\partial y\partial x} - y\frac{\partial}{\partial y} \\ & - yz\frac{\partial^2}{\partial y\partial z} - zx\frac{\partial^2}{\partial z\partial x} - zy\frac{\partial^2}{\partial z\partial y} - z\frac{\partial}{\partial z}\Bigg].\end{aligned} \tag{G.10}$$

Finally, comparing Eq. (G.7) with Eq. (G.10), we find the result of Eq. (5.136):

$$\begin{aligned}\hat{L}^2 &= \hat{r}^2\hat{p}^2 - \left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 + \hbar^2\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right) \\ &= \hat{r}^2\hat{p}^2 - \left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right)^2 + i\hbar\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right).\end{aligned} \tag{G.11}$$

Appendix H
Quantum states of the SU(2) oscillator

As explained in Chapter 7, the *n*th quantum state of a one-dimensional oscillator is obtained by applying the creation operator n times to the vacuum state. Let us use the simple ket notation $|n\rangle$ to denote the normalized result:

$$|n\rangle = N\,\left(a^{\dagger}\right)^{n}|0\rangle, \tag{H.1}$$

where n denotes the number of excitations and N is a normalizing factor. The corresponding bra is given by

$$\langle n| = N^{*}\ \langle 0|\,(a)^{n}. \tag{H.2}$$

The normalizer N is obtained from

$$\langle n|n\rangle = |N|^{2}\ \langle 0|\,(a)^{n}\left(a^{\dagger}\right)^{n}|0\rangle = 1. \tag{H.3}$$

This bracket can be determined from the commutation relationships of the creation and annihilation operators. We have

$$\begin{aligned} a\left(a^{\dagger}\right)^{n}|0\rangle &= n\left(a^{\dagger}\right)^{n-1}|0\rangle \\ (a)^{n}\left(a^{\dagger}\right)^{n}|0\rangle &= n!\,|0\rangle. \end{aligned} \tag{H.4}$$

The normalized eigenstates thus read

$$|n\rangle = \frac{1}{\sqrt{n!}}\left(a^{\dagger}\right)^{n}|0\rangle. \tag{H.5}$$

This state can be viewed as the creation of n particles. It is clear that the interchange of two creation operators does not change the state, so the quantum state is symmetric under particle exchange. The corresponding particles are therefore called *bosons*, and the state may be seen as the excitation of n bosons.

For a two-dimensional oscillator, two degrees of freedom can be excited. Following Chapter 8, we label these as the two spinor components $\{\alpha, \beta\}$. The only nonzero commutation relations are between the creation and annihilation operators of the same component; hence,

$$\left[a_{\alpha}, a_{\alpha}^{\dagger}\right] = \hbar\omega; \tag{H.6}$$

$$\left[a_{\beta}, a_{\beta}^{\dagger}\right] = \hbar\omega. \tag{H.7}$$

All commutators between operators for different components vanish. For example,

$$\left[a_{\alpha}, a_{\beta}^{\dagger}\right] = \frac{1}{2}\left[\left(a_{x} - ia_{y}\right), \left(a_{x}^{\dagger} - ia_{y}^{\dagger}\right)\right] = \frac{1}{2}\left(\left[a_{x}, a_{x}^{\dagger}\right] - \left[a_{y}, a_{y}^{\dagger}\right]\right) = 0. \tag{H.8}$$

As a result, the quantum states of SU(2) simply reduce to the product of two oscillator states. Let us denote the resulting ket function as $|n, m\rangle$, where n α-particles and m β-particles are created:

$$|n, m\rangle = \frac{1}{\sqrt{n!m!}} \left(a_\alpha^\dagger\right)^n \left(a_\beta^\dagger\right)^m |0\rangle. \tag{H.9}$$

Application of the Dirac operators to this state can be worked out easily. For example,

$$a_\alpha^\dagger |n, m\rangle = \frac{1}{\sqrt{n!m!}} \left(a_\alpha^\dagger\right)^{n+1} \left(a_\beta^\dagger\right)^m |0\rangle = \sqrt{n+1}\; |n+1, m\rangle; \tag{H.10}$$

$$a_\beta |n, m\rangle = \frac{1}{\sqrt{n!m!}} \left(a_\alpha^\dagger\right)^n a_\beta \left(a_\beta^\dagger\right)^m |0\rangle = \sqrt{m}\; |n, m-1\rangle. \tag{H.11}$$

In Chapter 8 we studied the spin operators acting on a basic spinor, with components $\{\alpha, \beta\}$:

$$\begin{aligned} \hat{S}_x &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\beta + a_\beta^\dagger a_\alpha\right); \\ \hat{S}_y &= \frac{\hbar}{2i}\left(a_\alpha^\dagger a_\beta - a_\beta^\dagger a_\alpha\right); \\ \hat{S}_z &= \frac{\hbar}{2}\left(a_\alpha^\dagger a_\alpha - a_\beta^\dagger a_\beta\right). \end{aligned} \tag{H.12}$$

These operators are complemented by the *trace* or *number* operator that appears in the Hamiltonian. We define

$$\hat{N} = \left(a_\alpha^\dagger a_\alpha + a_\beta^\dagger a_\beta\right). \tag{H.13}$$

This operator counts the total number of boson excitations. Using the results from Eqs. (H.10) and (H.11), we obtain

$$\hat{N}\, |n, m\rangle = a_\alpha^\dagger a_\alpha\, |n, m\rangle + a_\beta^\dagger a_\beta\, |n, m\rangle = (n+m)\; |n, m\rangle. \tag{H.14}$$

This expression indicates that all states with the same $n+m$ excitation total are degenerate. In addition, the $|n, m\rangle$ quantum states are also eigenfunctions of the $\hat{S}_z$ operator:

$$\hat{S}_z\, |n, m\rangle = \frac{\hbar}{2} a_\alpha^\dagger a_\alpha\, |n, m\rangle - \frac{\hbar}{2} a_\beta^\dagger a_\beta\, |n, m\rangle = \frac{\hbar}{2}(n-m)\; |n, m\rangle. \tag{H.15}$$

We can define raising and lowering operators exactly as in the case of angular momentum:

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y = \hbar\, a_\alpha^\dagger a_\beta; \tag{H.16}$$

$$\hat{S}_- = \hat{S}_x - i\hat{S}_y = \hbar\, a_\beta^\dagger a_\alpha. \tag{H.17}$$

These raise or lower the value of the z-component by one quantum. We have

$$\hat{S}_+\, |n, m\rangle = \hbar\sqrt{(n+1)\, m}\; |n+1, m-1\rangle; \tag{H.18}$$

$$\hat{S}_-\, |n, m\rangle = \hbar\sqrt{n\,(m+1)}\; |n-1, m+1\rangle. \tag{H.19}$$

Combining these expressions, we can now calculate the total spin associated with a given SU(2) multiplet. Exactly the same as total angular momentum, spin momentum may be expressed with the help of the component operators as follows:

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = \hat{S}_-\, \hat{S}_+ + \hat{S}_z^2 + \hbar\hat{S}_z = \hbar^2\, \frac{\hat{N}}{2}\left(\frac{\hat{N}}{2} + 1\right). \tag{H.20}$$

Hence, the square of the total spin associated with the n, m quantum states is given by

$$\langle n, m|\, \hat{S}^2\, |n, m\rangle = \hbar^2\; \frac{n+m}{2}\left(\frac{n+m}{2} + 1\right). \tag{H.21}$$

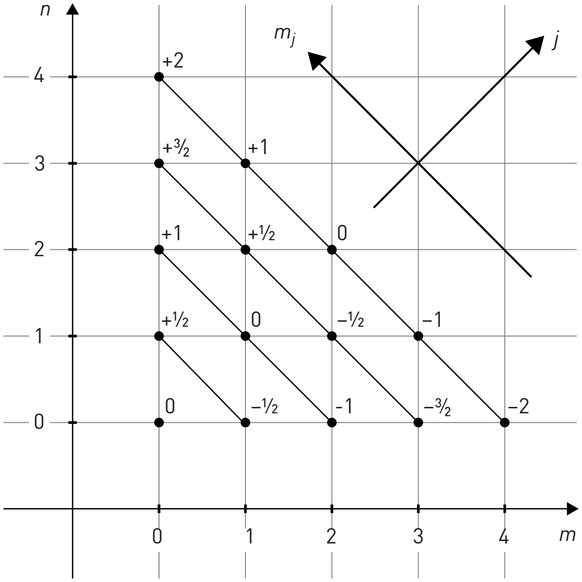


FIGURE H.1 Diagram of the $|n, m\rangle$ oscillator states of SU(2), with $m_j = 1/2\,(n - m)$ and $j = 1/2\,(n + m)$. The diagonal lines represent j-multiplets with $j = 0, 1/2, 1, 3/2, 2, \ldots$.

Since this result also corresponds to $\hbar^2\, j\big(j+1\big)$, the j quantum number and its m_j companion are given by

$$j = \frac{1}{2}\,(n + m); \tag{H.22}$$

$$m_j = \frac{1}{2}\,(n - m). \tag{H.23}$$

In Figure H.1, j-multiplets as a function of n and m are displayed.

Appendix I

Commutation relations

A complete derivation of the commutation relations between $\hat{\mathbf{L}}$, $\hat{\mathbf{M}}$, and $\hat{\mathscr{H}}$ is not easily found in the scientific literature. All too many books leave their readers mystified regarding how they might obtain the given results.[1] The reason for this is that the proofs, despite the relatively simple mathematics involved, are rather lengthy and tedious. This appendix incorporates all the derivations necessary to obtain the relevant commutation relations, which are listed in Table I.1 for easy reference.

Table I.1 Overview of the most important commutation relations.

1.	$\left[\hat r_i, \hat p_j\right]$	$= i\hbar\delta_{ij}$	2.	$\left[\hat r_i, \hat r_j\right]$	$= 0$
3.	$\left[\hat p_i, \hat p_j\right]$	$= 0$	4.	$\left[\hat L_i, \hat r_j\right]$	$= i\hbar\varepsilon_{ijk}\hat r_k$
5.	$\left[\hat L_i, \hat p_j\right]$	$= i\hbar\varepsilon_{ijk}\hat p_k$	6.	$\left[\hat L_i, \hat L_j\right]$	$= i\hbar\varepsilon_{ijk}\hat L_k$
7.	$\left[\hat p_i, \frac{\hat r_j}{r}\right]$	$= -i\hbar\left(\delta_{ij}\frac{1}{r} - \frac{\hat r_i\hat r_j}{r^3}\right)$	8.	$\left[\hat L_i, \frac{\hat r_j}{r}\right]$	$= i\hbar\varepsilon_{ijk}\frac{\hat r_k}{r}$
9.	$\left[\hat M_i, \hat L_j\right]$	$= i\hbar\varepsilon_{ijk}\hat M_k$	10.	$\left[\hat M_i, \hat M_j\right]$	$= i\hbar\left(-2m\hat{\mathscr{H}}\right)\varepsilon_{ijk}\hat L_k$
11.	$\left[\hat L_i, \hat{\mathscr{H}}\right]$	$= 0$	12.	$\left[\hat M_i, \hat{\mathscr{H}}\right]$	$= 0$

I.1 CANONICAL COMMUTATION RELATIONS

I.1.1 Commutation relations for $\left[\hat r_i, \hat p_j\right]$

Position and momentum are conjugate operators and obey Heisenberg's *canonical commutation relations*, as will be established in this section. Before we get down to business, we must remember that working with operators can be a tricky thing; mistakes are bound to slip into the calculations when we forget that operators always *act* on a function. In this light, the concept of a *test function* f might come in handy. Let us examine its use by deriving the commutator for the x-components of the $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ operators:

$$\left[x, \hat p_x\right] = \left[x, \frac{\hbar}{i}\frac{\partial}{\partial x}\right]. \tag{I.1}$$

---

[1] Notable exceptions are Greiner and Müller's *Quantum Mechanics: Symmetries* and Burkhardt and Leventhal's *Topics in Atomic Physics*, which we closely follow in the following derivations. See W. Greiner and B. Müller. "Dynamical Symmetries". In: *Quantum Mechanics: Symmetries*. Berlin: Springer-Verlag, 2001, pp. 477–496; and C. E. Burkhardt and J. J. Leventhal. *Topics in Atomic Physics*. New York: Springer, 2006. The latter does not delve into group theory, however.

Since the commutator of two operators is itself an operator, we let it act on our test function f:

$$\begin{aligned}\left[x,\hat{p}_x\right]f &= \left[x,\frac{\hbar}{i}\frac{\partial}{\partial x}\right]f \\ &= \frac{\hbar}{i}\left(x\frac{\partial}{\partial x}f-\frac{\partial}{\partial x}xf\right) \\ &= \frac{\hbar}{i}\left(x\frac{\partial f}{\partial x}-\frac{dx}{dx}f-x\frac{\partial f}{\partial x}\right)=i\hbar f.\end{aligned} \tag{I.2}$$

At this point, we can drop the test function, to obtain

$$\left[x,\hat{p}_x\right]=i\hbar. \tag{I.3}$$

Note that in this, and most of the following computations, we will no longer write down the test function explicitly, so be warned! By substitution of y and z for x, we can generalize Eq. (I.3) to

$$\left[\hat{r}_i,\hat{p}_i\right]=i\hbar, \tag{I.4}$$

where the index i stands for x, y, or z, and $\hat{r}_x=x$, $\hat{r}_y=y$, and $\hat{r}_z=z$. However,

$$\left[x,\hat{p}_y\right]=\left[x,\frac{\hbar}{i}\frac{\partial}{\partial y}\right]=\frac{\hbar}{i}\left(x\frac{\partial}{\partial y}-\frac{\partial}{\partial y}x\right)=\frac{\hbar}{i}\left(x\frac{\partial}{\partial y}-0-x\frac{\partial}{\partial y}\right)=0; \tag{I.5}$$

$$\left[x,\hat{p}_z\right]=\left[x,\frac{\hbar}{i}\frac{\partial}{\partial z}\right]=\frac{\hbar}{i}\left(x\frac{\partial}{\partial z}-\frac{\partial}{\partial z}x\right)=\frac{\hbar}{i}\left(x\frac{\partial}{\partial z}-0-x\frac{\partial}{\partial z}\right)=0. \tag{I.6}$$

Generalizing by two cyclic permutations of x, y, and z, we obtain

$$\left[\hat{r}_i,\hat{p}_j\right]=0, \qquad \text{for } i\neq j. \tag{I.7}$$

Combining the commutation relations in Eq. (I.4) and Eq. (I.7) then leads to the canonical commutation relations for position and momentum:

$$\left[\hat{r}_i,\hat{p}_j\right]=i\hbar\delta_{ij}. \tag{I.8}$$

I.1.2 Commutation relations for $\left[\hat{r}_i,\hat{r}_j\right]$ and $\left[\hat{p}_i,\hat{p}_j\right]$

We find analogously that the position operators $\hat{r}_i$ mutually commute:

$$\left[\hat{r}_i,\hat{r}_j\right]=0. \tag{I.9}$$

The linear momentum operators also commute among themselves:

$$\left[\hat{p}_i,\hat{p}_j\right]=0. \tag{I.10}$$

I.2 ANGULAR MOMENTUM COMMUTATION RELATIONS

I.2.1 Commutation relations for $\left[\hat{L}_i, \hat{r}_j\right]$

Let us now evaluate the commutators of the components of the angular momentum operator $\hat{\mathbf{L}}$ and the position operator $\hat{\mathbf{r}}$. To this end, we write

$$\begin{aligned}\left[\hat{L}_x, x\right] &= \left[y\hat{p}_z - z\hat{p}_y, x\right] = \left[y\hat{p}_z, x\right] - \left[z\hat{p}_y, x\right] \\ &= \left[y, x\right]\hat{p}_z + y\left[\hat{p}_z, x\right] - \left[z, x\right]\hat{p}_y - z\left[\hat{p}_y, x\right] \\ &= 0.\hat{p}_z + y.0 - 0.\hat{p}_y - z.0 = 0,\end{aligned} \tag{I.11}$$

$$\begin{aligned}\left[\hat{L}_x, y\right] &= \left[y\hat{p}_z - z\hat{p}_y, y\right] = \left[y\hat{p}_z, y\right] - \left[z\hat{p}_y, y\right] \\ &= \left[y, y\right]\hat{p}_z + y\left[\hat{p}_z, y\right] - \left[z, y\right]\hat{p}_y - z\left[\hat{p}_y, y\right] \\ &= 0.\hat{p}_z + y.0 - 0.\hat{p}_y + z.i\hbar = i\hbar z,\end{aligned} \tag{I.12}$$

$$\begin{aligned}\left[\hat{L}_x, z\right] &= \left[y\hat{p}_z - z\hat{p}_y, z\right] = \left[y\hat{p}_z, z\right] - \left[z\hat{p}_y, z\right] \\ &= \left[y, z\right]\hat{p}_z + y\left[\hat{p}_z, z\right] - \left[z, z\right]\hat{p}_y - z\left[\hat{p}_y, z\right] \\ &= 0.\hat{p}_z - y.i\hbar - 0.\hat{p}_y - z.0 = -i\hbar y,\end{aligned} \tag{I.13}$$

where we have applied Eq. (I.8) and Eq. (I.9), as well as the commutation relation[2]

$$\left[\hat{A}\hat{B}, \hat{C}\right] = \left[\hat{A}, \hat{C}\right]\hat{B} + \hat{A}\left[\hat{B}, \hat{C}\right]. \tag{I.15}$$

Generalizing Eqs. (I.11) and (I.12) via two successive cyclic permutations leads to

$$\left[\hat{L}_i, \hat{r}_j\right] = i\hbar\varepsilon_{ijk}\hat{r}_k, \tag{I.16}$$

where ε_{ijk} is the Levi-Civita antisymmetric tensor.

I.2.2 Commutation relations for $\left[\hat{L}_i, \hat{p}_j\right]$

The commutation relations between the components of the angular momentum operator $\hat{\mathbf{L}}$ and the linear momentum operator $\hat{\mathbf{p}}$ can be obtained analogously. We thus write

$$\begin{aligned}\left[\hat{L}_x, \hat{p}_x\right] &= \left[y\hat{p}_z - z\hat{p}_y, \hat{p}_x\right] = \left[y\hat{p}_z, \hat{p}_x\right] - \left[z\hat{p}_y, \hat{p}_x\right] \\ &= \left[y, \hat{p}_x\right]\hat{p}_z + y\left[\hat{p}_z, \hat{p}_x\right] - \left[z, \hat{p}_x\right]\hat{p}_y - z\left[\hat{p}_y, \hat{p}_x\right] \\ &= 0.\hat{p}_z + y.0 - 0.\hat{p}_y - z.0 = 0,\end{aligned} \tag{I.17}$$

$$\begin{aligned}\left[\hat{L}_x, \hat{p}_y\right] &= \left[y\hat{p}_z - z\hat{p}_y, \hat{p}_y\right] = \left[y\hat{p}_z, \hat{p}_y\right] - \left[z\hat{p}_y, \hat{p}_y\right] \\ &= \left[y, \hat{p}_y\right]\hat{p}_z + y\left[\hat{p}_z, \hat{p}_y\right] - \left[z, \hat{p}_y\right]\hat{p}_y - z\left[\hat{p}_y, \hat{p}_y\right] \\ &= i\hbar\hat{p}_z + y.0 - 0.\hat{p}_y - z.0 = i\hbar\hat{p}_z,\end{aligned} \tag{I.18}$$

---

[2] The commutation relation in Eq. (I.15) is proved as follows:

$$\begin{aligned}\left[\hat{A}\hat{B}, \hat{C}\right] &= \hat{A}\hat{B}\hat{C} - \hat{C}\hat{A}\hat{B} \\ &= \hat{A}\hat{B}\hat{C} - \hat{C}\hat{A}\hat{B} + \hat{A}\hat{C}\hat{B} - \hat{A}\hat{C}\hat{B} \\ &= \hat{A}\left(\hat{B}\hat{C} - \hat{C}\hat{B}\right) + \left(\hat{A}\hat{C} - \hat{C}\hat{A}\right)\hat{B} \\ &= \hat{A}\left[\hat{B}, \hat{C}\right] + \left[\hat{A}, \hat{C}\right]\hat{B}.\end{aligned} \tag{I.14}$$

$$
\begin{aligned}
\left[\hat{L}_x, \hat{p}_z\right] &= \left[y\hat{p}_z - z\hat{p}_y, \hat{p}_z\right] = \left[y\hat{p}_z, \hat{p}_z\right] - \left[z\hat{p}_y, \hat{p}_z\right] \\
&= \left[y, \hat{p}_z\right]\hat{p}_z + y\left[\hat{p}_z, \hat{p}_z\right] - \left[z, \hat{p}_z\right]\hat{p}_y - z\left[\hat{p}_y, \hat{p}_z\right] \\
&= 0.\hat{p}_y + y.0 - i\hbar\hat{p}_y - z.0 = -i\hbar\hat{p}_y,
\end{aligned}
\tag{I.19}
$$

where we have used Eqs. (I.8) and (I.10), and the commutation relation in Eq. (I.15). After two more successive cyclic permutations, we obtain an expression for the commutator of $\hat{\mathbf{L}}$ and $\hat{\mathbf{p}}$:

$$\left[\hat{L}_i, \hat{p}_j\right] = i\hbar\varepsilon_{ijk}\hat{p}_k, \tag{I.20}$$

where ε_{ijk} is the Levi-Civita antisymmetric tensor.

I.2.3 Commutation relations for $\left[\hat{L}_i, \hat{L}_j\right]$

The angular momentum commutation relations were derived in Chapter 5:

$$\left[\hat{L}_i, \hat{L}_j\right] = i\hbar\varepsilon_{ijk}\hat{L}_k. \tag{I.21}$$

For the sake of completeness, let us derive them once again using the canonical commutation relations of §I.1. We thus write

$$\left[\hat{L}_i, \hat{L}_j\right] = \left[\left(\hat{\mathbf{r}} \times \hat{\mathbf{p}}\right)_i, \left(\hat{\mathbf{r}} \times \hat{\mathbf{p}}\right)_j\right]. \tag{I.22}$$

The cross products in Eq. (I.22) can be written in tensor notation, following Eq. (K.21):

$$\left[\hat{L}_i, \hat{L}_j\right] = \left[\varepsilon_{ikl}\hat{r}_k\hat{p}_l, \varepsilon_{jmn}\hat{r}_m\hat{p}_n\right] = \varepsilon_{ikl}\varepsilon_{jmn}\left[\hat{r}_k\hat{p}_l, \hat{r}_m\hat{p}_n\right]. \tag{I.23}$$

Recalling the elementary commutator identities $\left[\hat{A}\hat{B}, \hat{C}\right] = \left[\hat{A}, \hat{C}\right]\hat{B} + \hat{A}\left[\hat{B}, \hat{C}\right]$ and $\left[\hat{A}, \hat{B}\hat{C}\right] = \left[\hat{A}, \hat{B}\right]\hat{C} + \hat{B}\left[\hat{A}, \hat{C}\right]$, we have

$$
\begin{aligned}
\left[\hat{r}_k\hat{p}_l, \hat{r}_m\hat{p}_n\right] &= \left[\hat{r}_k\hat{p}_l, \hat{r}_m\right]\hat{p}_n + \hat{r}_m\left[\hat{r}_k\hat{p}_l, \hat{p}_n\right] \\
&= \left[\hat{r}_k, \hat{r}_m\right]\hat{p}_l\hat{p}_n + \hat{r}_k\left[\hat{p}_l, \hat{r}_m\right]\hat{p}_n \\
&\quad + \hat{r}_m\left[\hat{r}_k, \hat{p}_n\right]\hat{p}_l + \hat{r}_m\hat{r}_k\left[\hat{p}_l, \hat{p}_n\right] \\
&= -i\hbar\delta_{lm}\hat{r}_k\hat{p}_n + i\hbar\delta_{kn}\hat{r}_m\hat{p}_l.
\end{aligned}
\tag{I.24}
$$

Substitution in Eq. (I.23) yields

$$
\begin{aligned}
\left[\hat{L}_i, \hat{L}_j\right] &= \varepsilon_{ikl}\varepsilon_{jmn}\left(-i\hbar\delta_{lm}\hat{r}_k\hat{p}_n + i\hbar\delta_{kn}\hat{r}_m\hat{p}_l\right) \\
&= -i\hbar\varepsilon_{ikl}\varepsilon_{jln}\hat{r}_k\hat{p}_n + i\hbar\varepsilon_{ikl}\varepsilon_{jmk}\hat{r}_m\hat{p}_l.
\end{aligned}
\tag{I.25}
$$

The products of Levi-Civita symbols ($\varepsilon_{ikl}\varepsilon_{jln}$ and $\varepsilon_{ikl}\varepsilon_{jmk}$) can be written in terms of Kronecker deltas with the help of the contracted epsilon identities in Eq. (K.5):

$$\varepsilon_{ikl}\varepsilon_{jln} = \varepsilon_{lik}\varepsilon_{lnj} = \delta_{in}\delta_{kj} - \delta_{ij}\delta_{kn}, \tag{I.26}$$

$$\varepsilon_{ikl}\varepsilon_{jmk} = \varepsilon_{kli}\varepsilon_{kjm} = \delta_{lj}\delta_{im} - \delta_{lm}\delta_{ij}, \tag{I.27}$$

yielding

$$
\begin{aligned}
\left[\hat{L}_i, \hat{L}_j\right] &= -i\hbar\left(\delta_{in}\delta_{kj} - \delta_{ij}\delta_{kn}\right)\hat{r}_k\hat{p}_n + i\hbar\left(\delta_{lj}\delta_{im} - \delta_{lm}\delta_{ij}\right)\hat{r}_m\hat{p}_l \\
&= -i\hbar\left(\hat{r}_j\hat{p}_i - \delta_{ij}\hat{r}_k\hat{p}_k\right) + i\hbar\left(\hat{r}_i\hat{p}_j - \delta_{ij}\hat{r}_l\hat{p}_l\right) \\
&= i\hbar\left(\hat{r}_i\hat{p}_j - \hat{r}_j\hat{p}_i\right).
\end{aligned}
\tag{I.28}
$$

We want to prove that $\left[\hat{L}_i, \hat{L}_j\right] = i\hbar\varepsilon_{ijk}\hat{L}_k$. Rewriting $i\hbar\varepsilon_{ijk}\hat{L}_k$ as

$$i\hbar\varepsilon_{ijk}\hat{L}_k = i\hbar\varepsilon_{ijk}\left(\hat{\mathbf{r}} \times \hat{\mathbf{p}}\right)_k = i\hbar\varepsilon_{ijk}\varepsilon_{klm}\hat{r}_l\hat{p}_m, \tag{I.29}$$

and using the contracted epsilon identity from Eq. (K.5) to write

$$\varepsilon_{ijk}\varepsilon_{klm} = \varepsilon_{kij}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \tag{I.30}$$

finally yields

$$i\hbar\varepsilon_{ijk}\hat{L}_k = i\hbar\left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)\hat{r}_l\hat{p}_m = i\hbar\left(\hat{r}_i\hat{p}_j - \hat{r}_j\hat{p}_i\right), \tag{I.31}$$

which is identical to Eq. (I.28) and finalizes our proof.

I.3 AUXILIARY COMMUTATORS

Before determining $\left[\hat{M}_i, \hat{L}_j\right]$ and $\left[\hat{M}_i, \hat{M}_j\right]$, let us first derive a number of *auxiliary commutators.*

I.3.1 Commutation relations for $\left[\hat{p}_i, \hat{r}_j/r\right]$

The simplest way to obtain a general expression for the commutator $\left[\hat{p}_i, \hat{r}_j/r\right]$ is to first compute it explicitly for its components in a number of cases, and then generalize. We therefore start by deriving an expression for $\left[\hat{p}_i, \hat{r}_j/r\right]$ when $i \neq j$ and then move on to the case where $i = j$.

Case 1: The commutator $\left[\hat{p}_i, \hat{r}_j/r\right]$ for $i \neq j$

Let us derive, by way of example, the commutator $\left[\hat{p}_x, \frac{y}{r}\right]$. We have

$$\begin{aligned}\left[\hat{p}_x, \frac{y}{r}\right] &= \hat{p}_x\frac{y}{r} - \frac{y}{r}\hat{p}_x \\ &= -i\hbar\frac{\partial}{\partial x}\left(\frac{y}{r}\right) - i\hbar\frac{y}{r}\frac{\partial}{\partial x} + i\hbar\frac{y}{r}\frac{\partial}{\partial x} \\ &= -i\hbar\frac{\partial}{\partial x}\left(yr^{-1}\right).\end{aligned} \tag{I.32}$$

The partial derivative in the previous equation can then be written as

$$\begin{aligned}\frac{\partial}{\partial x}\left(yr^{-1}\right) &= \frac{1}{r}\frac{\partial y}{\partial x} + y\frac{\partial r^{-1}}{\partial x} \\ &= 0 + y\frac{\partial\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}}{\partial\left(x^2+y^2+z^2\right)}\frac{\partial\left(x^2+y^2+z^2\right)}{\partial x} \\ &= -\frac{1}{2}y\frac{1}{r^3}2x = -\frac{xy}{r^3}.\end{aligned} \tag{I.33}$$

Substituting this result in Eq. (I.32) leads to

$$\left[\hat{p}_x, \frac{y}{r}\right] = i\hbar\frac{xy}{r^3}. \tag{I.34}$$

This can be easily generalized to

$$\left[\hat{p}_i, \frac{\hat{r}_j}{r}\right] = i\hbar\frac{\hat{r}_i\hat{r}_j}{r^3}, \qquad \forall i \neq j. \tag{I.35}$$

Case 2: The commutator $\left[\hat{p}_i, \hat{r}_j/r\right]$ for $i = j$

In a next step, let us determine the commutator $\left[\hat{p}_x, x/r\right]$:

$$\begin{aligned}\left[\hat{p}_x, \frac{x}{r}\right] &= \hat{p}_x\frac{x}{r} - \frac{x}{r}\hat{p}_x \\ &= -i\hbar\frac{\partial}{\partial x}\left(\frac{x}{r}\right) - i\hbar\frac{x}{r}\frac{\partial}{\partial x} + i\hbar\frac{x}{r}\frac{\partial}{\partial x} \\ &= -i\hbar\frac{\partial}{\partial x}\left(xr^{-1}\right).\end{aligned} \tag{I.36}$$

We can rewrite the partial derivative as

$$\begin{aligned}\frac{\partial}{\partial x}\left(xr^{-1}\right) &= \frac{1}{r}\frac{\partial x}{\partial x}+x\frac{\partial r^{-1}}{\partial x}\\ &= \frac{1}{r}+x\frac{\partial\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}}{\partial\left(x^2+y^2+z^2\right)}\frac{\partial\left(x^2+y^2+z^2\right)}{\partial x}\\ &= \frac{1}{r}-\frac{1}{2}x\frac{1}{r^3}2x=\frac{1}{r}-\frac{x^2}{r^3},\end{aligned} \tag{I.37}$$

which leads to

$$\left[\hat{p}_x,\frac{x}{r}\right]=-i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right). \tag{I.38}$$

Similarly, we obtain

$$\left[\hat{p}_y,\frac{y}{r}\right]=-i\hbar\left(\frac{1}{r}-\frac{y^2}{r^3}\right); \tag{I.39}$$

$$\left[\hat{p}_z,\frac{z}{r}\right]=-i\hbar\left(\frac{1}{r}-\frac{z^2}{r^3}\right). \tag{I.40}$$

Generalization

Combining the results of both cases leads to the following expression for the commutator $\left[\hat{p}_i,\frac{\hat{r}_j}{r}\right]$:

$$\left[\hat{p}_i,\frac{\hat{r}_j}{r}\right]=-i\hbar\left(\delta_{ij}\frac{1}{r}-\frac{\hat{r}_i\hat{r}_j}{r^3}\right). \tag{I.41}$$

I.3.2 Commutation relations for $\left[\hat{L}_i,\hat{r}_j/r\right]$

As in the preceding section, it will prove convenient to first evaluate this commutator for its components in a number of cases. Let us proceed by deriving an expression for $\left[\hat{L}_i,\hat{r}_j/r\right]$ when $i\neq j$ and then move on to the case where $i=j$.

Case 1: The commutator $\left[\hat{L}_i,\hat{r}_j/r\right]$ for $i\neq j$

Let us, by way of example, derive the commutator $\left[\hat{L}_x,y/r\right]$. We have

$$\begin{aligned}\left[\hat{L}_x,\frac{y}{r}\right] &= \hat{L}_x\frac{y}{r}-\frac{y}{r}\hat{L}_x\\ &= \left(y\hat{p}_z-z\hat{p}_y\right)\frac{y}{r}-\frac{y}{r}\left(y\hat{p}_z-z\hat{p}_y\right)\\ &= y\hat{p}_z\frac{y}{r}-z\hat{p}_y\frac{y}{r}-\frac{y^2}{r}\hat{p}_z+\frac{yz}{r}\hat{p}_y.\end{aligned} \tag{I.42}$$

The first term can be written as follows:

$$\begin{aligned}y\hat{p}_z\frac{y}{r} &= -i\hbar y\frac{\partial}{\partial z}\frac{y}{r}-i\hbar y\frac{y}{r}\frac{\partial}{\partial z}\\ &= -i\hbar y\frac{\partial}{\partial z}\frac{y}{r}+\frac{y^2}{r}\hat{p}_z.\end{aligned} \tag{I.43}$$

Analogously, for the second term of Eq. (I.42),

$$\begin{aligned}z\hat{p}_y\frac{y}{r} &= -i\hbar z\frac{\partial}{\partial y}\frac{y}{r}-i\hbar z\frac{y}{r}\frac{\partial}{\partial y}\\ &= -i\hbar z\frac{\partial}{\partial y}\frac{y}{r}+\frac{yz}{r}\hat{p}_y.\end{aligned} \tag{I.44}$$

Substitution of these results in Eq. (I.42) yields

$$\left[\hat{L}_x, \frac{y}{r}\right] = -i\hbar y \frac{\partial}{\partial z}\frac{y}{r} + i\hbar z \frac{\partial}{\partial y}\frac{y}{r}. \tag{I.45}$$

The first partial derivative can be written as

$$\begin{aligned} \frac{\partial}{\partial z}\frac{y}{r} &= \frac{1}{r}\frac{\partial y}{\partial z} + y\frac{\partial r^{-1}}{\partial z} \\ &= 0 + y\frac{\partial\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}}{\partial\left(x^2+y^2+z^2\right)}\frac{\partial\left(x^2+y^2+z^2\right)}{\partial z} \\ &= -\frac{1}{2}y\frac{1}{r^3}2z = -\frac{yz}{r^3}. \end{aligned} \tag{I.46}$$

Similarly, for the second derivative,

$$\begin{aligned} \frac{\partial}{\partial y}\frac{y}{r} &= \frac{1}{r}\frac{\partial y}{\partial y} + y\frac{\partial r^{-1}}{\partial y} \\ &= \frac{1}{r} + y\frac{\partial\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}}{\partial\left(x^2+y^2+z^2\right)}\frac{\partial\left(x^2+y^2+z^2\right)}{\partial y} \\ &= \frac{1}{r} - \frac{1}{2}y\frac{1}{r^3}2y = \frac{1}{r} - \frac{y^2}{r^3}. \end{aligned} \tag{I.47}$$

The commutator in Eq. (I.45) then reduces to

$$\begin{aligned} \left[\hat{L}_x, \frac{y}{r}\right] &= i\hbar\frac{y^2 z}{r^3} + i\hbar z\left(\frac{1}{r} - \frac{y^2}{r^3}\right) \\ &= i\hbar\frac{y^2 z}{r^3} + i\hbar\frac{z}{r} - i\hbar\frac{y^2 z}{r^3} \\ \left[\hat{L}_x, \frac{y}{r}\right] &= i\hbar\frac{z}{r}. \end{aligned} \tag{I.48}$$

Two successive cyclic permutations of x, y, and z yield two more commutators:

$$\left[\hat{L}_y, \frac{z}{r}\right] = i\hbar\frac{x}{r}, \qquad \left[\hat{L}_z, \frac{x}{r}\right] = i\hbar\frac{y}{r}. \tag{I.49}$$

Case 2: The commutator $\left[\hat{L}_i, \hat{r}_j/r\right]$ for $i = j$

Next, let us determine the commutator $\left[\hat{L}_x, x/r\right]$:

$$\begin{aligned} \left[\hat{L}_x, \frac{x}{r}\right] &= \hat{L}_x\frac{x}{r} - \frac{x}{r}\hat{L}_x \\ &= \left(y\hat{p}_z - z\hat{p}_y\right)\frac{x}{r} - \frac{x}{r}\left(y\hat{p}_z - z\hat{p}_y\right) \\ &= y\hat{p}_z\frac{x}{r} - z\hat{p}_y\frac{x}{r} - \frac{xy}{r}\hat{p}_z + \frac{xz}{r}\hat{p}_y. \end{aligned} \tag{I.50}$$

After rewriting the first and second terms as

$$\begin{aligned} y\hat{p}_z\frac{x}{r} &= -i\hbar y\frac{\partial}{\partial z}\frac{x}{r} - i\hbar y\frac{x}{r}\frac{\partial}{\partial z} \\ &= -i\hbar y\frac{\partial}{\partial z}\frac{x}{r} + \frac{xy}{r}\hat{p}_z, \end{aligned} \tag{I.51}$$

and

$$\begin{aligned} z\hat{p}_y\frac{x}{r} &= -i\hbar z\frac{\partial}{\partial y}\frac{x}{r} - i\hbar z\frac{x}{r}\frac{\partial}{\partial y} \\ &= -i\hbar z\frac{\partial}{\partial y}\frac{x}{r} + \frac{xz}{r}\hat{p}_y, \end{aligned} \tag{I.52}$$

we obtain

$$\left[\hat{L}_x, \frac{x}{r}\right] = -i\hbar y \frac{\partial}{\partial z}\frac{x}{r} + i\hbar z \frac{\partial}{\partial y}\frac{x}{r}. \tag{I.53}$$

Rewriting the two partial derivatives as

$$\begin{aligned} \frac{\partial}{\partial z}\frac{x}{r} &= \frac{1}{r}\frac{\partial x}{\partial z} + x\frac{\partial r^{-1}}{\partial z} \\ &= 0 + x\frac{\partial\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}}{\partial\left(x^2+y^2+z^2\right)}\frac{\partial\left(x^2+y^2+z^2\right)}{\partial z} \\ &= -\frac{1}{2}x\frac{1}{r^3}2z = -\frac{xz}{r^3} \end{aligned} \tag{I.54}$$

and

$$\begin{aligned} \frac{\partial}{\partial y}\frac{x}{r} &= \frac{1}{r}\frac{\partial x}{\partial y} + x\frac{\partial r^{-1}}{\partial y} \\ &= 0 + x\frac{\partial\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}}{\partial\left(x^2+y^2+z^2\right)}\frac{\partial\left(x^2+y^2+z^2\right)}{\partial y} \\ &= -\frac{1}{2}x\frac{1}{r^3}2y = -\frac{xy}{r^3}, \end{aligned} \tag{I.55}$$

finally leads to

$$\left[\hat{L}_x, \frac{x}{r}\right] = i\hbar\frac{xyz}{r^3} - i\hbar\frac{xyz}{r^3} = 0. \tag{I.56}$$

Generalization

The remaining commutation relations can be determined by cyclic permutation of x, y, and z, and results in the generalized commutation relation

$$\left[\hat{L}_i, \frac{\hat{r}_j}{r}\right] = i\hbar\varepsilon_{ijk}\frac{\hat{r}_k}{r}, \tag{I.57}$$

where ε_{ijk} is the Levi-Civita antisymmetric tensor.

I.4 REFORMULATION OF THE LRL VECTOR

When evaluating the commutators $\left[\hat{M}_i, \hat{L}_j\right]$ and $\left[\hat{M}_i, \hat{M}_j\right]$, it is convenient to express the Laplace-Runge-Lenz (LRL) vector in a slightly different form. To this aim, recall that the operator form of the LRL vector was defined as

$$\hat{\mathbf{M}} = \frac{1}{2}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - \hat{\mathbf{L}}\times\hat{\mathbf{p}}\right) - mk\frac{\hat{\mathbf{r}}}{r}. \tag{I.58}$$

Given Eq. (K.21), the ith component of the cross product $\hat{\mathbf{L}}\times\hat{\mathbf{p}}$ can be written in tensor notation:

$$\left(\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)_i = \sum_{jk}\varepsilon_{ijk}\hat{L}_j\hat{p}_k. \tag{I.59}$$

From the commutation relations $\left[\hat{L}_j, \hat{p}_k\right] = i\hbar\varepsilon_{jkl}\hat{p}_l$, as derived in Eq. (I.20), we can write

$$\left(\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)_i = \sum_{jk}\varepsilon_{ijk}\hat{p}_k\hat{L}_j + i\hbar\sum_{jkl}\varepsilon_{ijk}\varepsilon_{jkl}\hat{p}_l. \tag{I.60}$$

The product of Levi-Civita symbols $\varepsilon_{ijk}\varepsilon_{jkl}$ in Eq. (I.60) can be written in terms of Kronecker deltas with the help of the contracted epsilon identity in Eq. (K.6):

$$\varepsilon_{ijk}\varepsilon_{jkl} = \varepsilon_{jki}\varepsilon_{jkl} = 2\delta_{il}, \tag{I.61}$$

yielding

$$\begin{aligned}\left(\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)_i &= -\sum_{jk}\varepsilon_{ikj}\hat{p}_k\hat{L}_j + 2i\hbar\sum_l \delta_{il}\hat{p}_l \\ &= -\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i + 2i\hbar\hat{p}_i\end{aligned}\tag{I.62}$$

or

$$\hat{\mathbf{L}}\times\hat{\mathbf{p}} = -\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right) + 2i\hbar\hat{\mathbf{p}}.\tag{I.63}$$

Hence, the LRL vector from Eq. (I.58) can be reformulated as

$$\begin{aligned}\hat{\mathbf{M}} &= \frac{1}{2}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}+\hat{\mathbf{p}}\times\hat{\mathbf{L}}-2i\hbar\hat{\mathbf{p}}\right) - mk\frac{\hat{\mathbf{r}}}{r} \\ &= \hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}} - mk\frac{\hat{\mathbf{r}}}{r}.\end{aligned}\tag{I.64}$$

I.5 COMMUTATORS WITH THE LRL OPERATOR

I.5.1 Commutation relations for $\left[\hat{M}_i,\hat{L}_j\right]$

As in the preceding cases, let us compute the commutator $\left[\hat{M}_i,\hat{L}_j\right]$ explicitly for its components in a number of cases. We start again by deriving an expression for $\left[\hat{M}_i,\hat{L}_j\right]$ when $i\neq j$ and then move on to the case where $i=j$.

Case 1: The commutator $\left[\hat{M}_i,\hat{L}_j\right]$ for $i\neq j$

Let us compute the commutator $\left[\hat{M}_x,\hat{L}_y\right]$. We have, using the new form of the LRL vector obtained in Eq. (I.64),

$$\begin{aligned}\left[\hat{M}_x,\hat{L}_y\right] &= \left[\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)_x - mk\frac{x}{r},\hat{L}_y\right] \\ &= \left[\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_x,\hat{L}_y\right] - \left[i\hbar\hat{p}_x,\hat{L}_y\right] - \left[mk\frac{x}{r},\hat{L}_y\right] \\ &= \left[\hat{p}_y\hat{L}_z-\hat{p}_z\hat{L}_y,\hat{L}_y\right] + i\hbar\left[\hat{L}_y,\hat{p}_x\right] + mk\left[\hat{L}_y,\frac{x}{r}\right].\end{aligned}\tag{I.65}$$

We then get

$$\begin{aligned}\left[\hat{M}_x,\hat{L}_y\right] &= \left(\hat{p}_y\hat{L}_z\hat{L}_y - \hat{p}_z\hat{L}_y\hat{L}_y - \hat{L}_y\hat{p}_y\hat{L}_z + \hat{L}_y\hat{p}_z\hat{L}_y\right) \\ &\quad + i\hbar\left[\hat{L}_y,\hat{p}_x\right] + mk\left[\hat{L}_y,\frac{x}{r}\right].\end{aligned}\tag{I.66}$$

Using the commutation relations $\left[L_i,L_j\right] = i\hbar\varepsilon_{ijk}L_k$ and $\left[\hat{L}_i,\hat{p}_j\right] = i\hbar\varepsilon_{ijk}\hat{p}_k$, as derived in §I.2, we can rewrite the term $\hat{p}_y\hat{L}_z\hat{L}_y$ as

$$\begin{aligned}\hat{p}_y\left(\hat{L}_z\hat{L}_y\right) &= \hat{p}_y\left(\hat{L}_y\hat{L}_z - i\hbar\hat{L}_x\right) \\ &= \hat{p}_y\hat{L}_y\hat{L}_z - i\hbar\hat{p}_y\hat{L}_x \\ &= \hat{L}_y\hat{p}_y\hat{L}_z - i\hbar\hat{p}_y\hat{L}_x,\end{aligned}\tag{I.67}$$

and the term $\hat{p}_z\hat{L}_y\hat{L}_y$ as

$$\begin{aligned}\left(\hat{p}_z\hat{L}_y\right)\hat{L}_y &= \left(\hat{L}_y\hat{p}_z - i\hbar\hat{p}_x\right)\hat{L}_y \\ &= \hat{L}_y\hat{p}_z\hat{L}_y - i\hbar\hat{p}_x\hat{L}_y.\end{aligned}\tag{I.68}$$

After substituting these results in the first term of Eq. (I.66), and canceling, we obtain

$$\left[\hat{M}_x,\hat{L}_y\right]=i\hbar\left(\hat{p}_x\hat{L}_y-\hat{p}_y\hat{L}_x\right)+i\hbar\left[\hat{L}_y,\hat{p}_x\right]+mk\left[\hat{L}_y,\frac{x}{r}\right]. \tag{I.69}$$

Using the commutation relations in Eq. (I.20) and Eq. (I.57), we can rewrite the second and third terms as

$$\left[\hat{L}_y,\hat{p}_x\right]=-i\hbar\hat{p}_z \tag{I.70}$$

and

$$\left[\hat{L}_y,\frac{x}{r}\right]=-i\hbar\frac{z}{r}. \tag{I.71}$$

We thus obtain

$$\begin{aligned}\left[\hat{M}_x,\hat{L}_y\right]&=i\hbar\left(\hat{p}_x\hat{L}_y-\hat{p}_y\hat{L}_x\right)+i\hbar\left(-i\hbar\hat{p}_z\right)-i\hbar mk\frac{z}{r}\\&=i\hbar\left[\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)_z-mk\frac{z}{r}\right]\\\left[\hat{M}_x,\hat{L}_y\right]&=i\hbar\hat{M}_z.\end{aligned} \tag{I.72}$$

Case 2: The commutator $\left[\hat{M}_i,\hat{L}_j\right]$ for $i=j$

In complete analogy, let us now derive the commutator $\left[\hat{M}_x,\hat{L}_x\right]$:

$$\begin{aligned}\left[\hat{M}_x,\hat{L}_x\right]&=\left[\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)_x-mk\frac{x}{r},\hat{L}_x\right]\\&=\left[\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_x,\hat{L}_x\right]-\left[i\hbar\hat{p}_x,\hat{L}_x\right]-\left[mk\frac{x}{r},\hat{L}_x\right]\\&=\left[\hat{p}_y\hat{L}_z-\hat{p}_z\hat{L}_y,\hat{L}_x\right]+i\hbar\left[\hat{L}_x,\hat{p}_x\right]+mk\left[\hat{L}_x,\frac{x}{r}\right].\end{aligned} \tag{I.73}$$

That is,

$$\begin{aligned}\left[\hat{M}_x,\hat{L}_x\right]=&\left(\hat{p}_y\hat{L}_z\hat{L}_x-\hat{p}_z\hat{L}_y\hat{L}_x-\hat{L}_x\hat{p}_y\hat{L}_z+\hat{L}_x\hat{p}_z\hat{L}_y\right)\\&+i\hbar\left[\hat{L}_x,\hat{p}_x\right]+mk\left[\hat{L}_x,\frac{x}{r}\right].\end{aligned} \tag{I.74}$$

Following Eq. (I.20) and Eq. (I.21), we can rewrite the terms $\hat{p}_y\hat{L}_z\hat{L}_x$ and $\hat{p}_z\hat{L}_y\hat{L}_x$ as

$$\begin{aligned}\hat{p}_y\hat{L}_z\hat{L}_x&=\hat{p}_y\hat{L}_x\hat{L}_z+i\hbar\hat{p}_y\hat{L}_y\\&=\hat{L}_x\hat{p}_y\hat{L}_z-i\hbar\hat{p}_z\hat{L}_z+i\hbar\hat{p}_y\hat{L}_y\end{aligned} \tag{I.75}$$

and

$$\begin{aligned}\hat{p}_z\hat{L}_y\hat{L}_x&=\hat{p}_z\hat{L}_x\hat{L}_y-i\hbar\hat{p}_z\hat{L}_z\\&=\hat{L}_x\hat{p}_z\hat{L}_y+i\hbar\hat{p}_y\hat{L}_y-i\hbar\hat{p}_z\hat{L}_z.\end{aligned} \tag{I.76}$$

Substituting these results in the first term of Eq. (I.74), and canceling yields

$$\hat{p}_y\hat{L}_z\hat{L}_x-\hat{p}_z\hat{L}_y\hat{L}_x-\hat{L}_x\hat{p}_y\hat{L}_z+\hat{L}_x\hat{p}_z\hat{L}_y=0, \tag{I.77}$$

so that

$$\left[\hat{M}_x,\hat{L}_x\right]=i\hbar\left[\hat{L}_x,\hat{p}_x\right)+mk\left[\hat{L}_x,\frac{x}{r}\right]. \tag{I.78}$$

Following Eq. (I.20) and Eq. (I.57), $\left[\hat{L}_x,\hat{p}_x\right]=0$ and $\left[\hat{L}_x,\frac{x}{r}\right]=0$. We thus obtain

$$\left[\hat{M}_x,\hat{L}_x\right]=0. \tag{I.79}$$

Generalization

The remaining commutation relations can be computed by cyclic permutation of x, y, and z, and they result in the generalized commutation relation

$$\left[\hat{M}_i, \hat{L}_j\right] = i\hbar\varepsilon_{ijk}\hat{M}_k, \tag{I.80}$$

where ε_{ijk} is the Levi-Civita antisymmetric tensor.

I.5.2 Commutation relations for $\left[\hat{\boldsymbol{M}}_i, \hat{\boldsymbol{M}}_j\right]$

Although frequent recourse will be needed to the commutation relations $\left[\hat{r}_i, \hat{r}_j\right] = \left[\hat{p}_i, \hat{p}_j\right] = 0$ and $\left[\hat{L}_i, \hat{p}_j\right] = i\hbar\varepsilon_{ijk}\hat{p}_k$, obtained in §I.1 and §I.2, no explicit mention will be made of their application in the following derivation.

Case 1: The commutator $\left[\hat{\boldsymbol{M}}_i, \hat{\boldsymbol{M}}_j\right]$ for $\boldsymbol{i \neq j}$

Let us evaluate, by way of illustration, the commutator $\left[\hat{M}_x, \hat{M}_y\right]$:

$$\begin{aligned}
\left[\hat{M}_x, \hat{M}_y\right] &= \hat{M}_x\hat{M}_y - \hat{M}_y\hat{M}_x \\
&= \left(\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)_x - mk\frac{x}{r}\right)\left(\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)_y - mk\frac{y}{r}\right) \\
&\quad - \left(\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)_y - mk\frac{y}{r}\right)\left(\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)_x - mk\frac{x}{r}\right) \qquad \text{(I.81)} \\
&= \left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y - i\hbar\hat{p}_x - mk\frac{x}{r}\right)\left(\hat{p}_z\hat{L}_x - \hat{p}_x\hat{L}_z - i\hbar\hat{p}_y - mk\frac{y}{r}\right) \\
&\quad - \left(\hat{p}_z\hat{L}_x - \hat{p}_x\hat{L}_z - i\hbar\hat{p}_y - mk\frac{y}{r}\right)\left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y - i\hbar\hat{p}_x - mk\frac{x}{r}\right),
\end{aligned}$$

where we have used the alternative form of the LRL operator, as derived in §I.4. Working out the brackets yields

$$\begin{aligned}
\left[\hat{M}_x, \hat{M}_y\right] &= \left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y - i\hbar\hat{p}_x\right)\left(\hat{p}_z\hat{L}_x - \hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\right) && \text{(I.82)} \\
&\quad - \left(\hat{p}_z\hat{L}_x - \hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\right)\left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y - i\hbar\hat{p}_x\right) && \text{(I.83)} \\
&\quad - mk\left(\left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y - i\hbar\hat{p}_x\right)\frac{y}{r} + \frac{x}{r}\left(\hat{p}_z\hat{L}_x - \hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\right)\right) && \text{(I.84)} \\
&\quad + mk\left(\left(\hat{p}_z\hat{L}_x - \hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\right)\frac{x}{r} + \frac{y}{r}\left(\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y - i\hbar\hat{p}_x\right)\right) && \text{(I.85)} \\
&\quad + m^2k^2\frac{xy}{r^2} - m^2k^2\frac{yx}{r^2}, && \text{(I.86)}
\end{aligned}$$

where the terms in Eq. (I.86) are seen to cancel. Performing the multiplication in Eq. (I.82) yields nine terms:

$$\begin{aligned}
&\hat{p}_y\hat{L}_z\hat{p}_z\hat{L}_x - \hat{p}_y\hat{L}_z\hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\hat{L}_z\hat{p}_y \\
&\quad - \hat{p}_z\hat{L}_y\hat{p}_z\hat{L}_x + \hat{p}_z\hat{L}_y\hat{p}_x\hat{L}_z + i\hbar\hat{p}_z\hat{L}_y\hat{p}_y \\
&\quad - i\hbar\hat{p}_x\hat{p}_z\hat{L}_x + i\hbar\hat{p}_x\hat{p}_x\hat{L}_z - \hbar^2\hat{p}_x\hat{p}_y.
\end{aligned} \tag{I.87}$$

After commuting the third and the eighth terms that is:

$$\begin{aligned}
i\hbar\hat{p}_y\hat{L}_z\hat{p}_y &= i\hbar\hat{p}_y\left(\hat{p}_y\hat{L}_z - i\hbar\hat{p}_x\right) \\
&= i\hbar\hat{p}_y\hat{p}_y\hat{L}_z + \hbar^2\hat{p}_y\hat{p}_x;
\end{aligned} \tag{I.88}$$

and

$$\begin{aligned} i\hbar\hat{p}_x\hat{p}_x\hat{L}_z &= i\hbar\hat{p}_x\left(\hat{L}_z\hat{p}_x - i\hbar\hat{p}_y\right) \\ &= i\hbar\hat{p}_x\hat{L}_z\hat{p}_x + \hbar^2\hat{p}_x\hat{p}_y, \end{aligned} \tag{I.89}$$

and rearranging the sixth, seventh, and ninth terms, we obtain

$$\begin{aligned} &\hat{p}_y\hat{L}_z\hat{p}_z\hat{L}_x - \hat{p}_y\hat{L}_z\hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\hat{p}_y\hat{L}_z \\ &\quad - \hat{p}_z\hat{L}_y\hat{p}_z\hat{L}_x + \hat{p}_z\hat{L}_y\hat{p}_x\hat{L}_z + i\hbar\hat{p}_y\hat{p}_z\hat{L}_y \\ &\quad - i\hbar\hat{p}_z\hat{L}_x\hat{p}_x + i\hbar\hat{p}_x\hat{L}_z\hat{p}_x - \hbar^2\hat{p}_y\hat{p}_x. \end{aligned} \tag{I.90}$$

The multiplication in Eq. (I.83) yields nine more terms:

$$\begin{aligned} &\hat{p}_z\hat{L}_x\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_x\hat{p}_z\hat{L}_y - i\hbar\hat{p}_z\hat{L}_x\hat{p}_x \\ &\quad - \hat{p}_x\hat{L}_z\hat{p}_y\hat{L}_z + \hat{p}_x\hat{L}_z\hat{p}_z\hat{L}_y + i\hbar\hat{p}_x\hat{L}_z\hat{p}_x \\ &\quad - i\hbar\hat{p}_y\hat{p}_y\hat{L}_z + i\hbar\hat{p}_y\hat{p}_z\hat{L}_y + i^2\hbar^2\hat{p}_y\hat{p}_x. \end{aligned} \tag{I.91}$$

Subtraction of Eq. (I.91) from Eq. (I.90) leaves, after canceling,

$$\begin{aligned} &\hat{p}_y\hat{L}_z\hat{p}_z\hat{L}_x - \hat{p}_y\hat{L}_z\hat{p}_x\hat{L}_z - \hat{p}_z\hat{L}_y\hat{p}_z\hat{L}_x + \hat{p}_z\hat{L}_y\hat{p}_x\hat{L}_z \\ &- \hat{p}_z\hat{L}_x\hat{p}_y\hat{L}_z + \hat{p}_z\hat{L}_x\hat{p}_z\hat{L}_y + \hat{p}_x\hat{L}_z\hat{p}_y\hat{L}_z - \hat{p}_x\hat{L}_z\hat{p}_z\hat{L}_y. \end{aligned} \tag{I.92}$$

Let us simplify this expression by rewriting the first, fourth, sixth, and seventh terms in the following form:

$$\begin{aligned} \hat{p}_y\hat{L}_z\hat{p}_z\hat{L}_x &= \hat{p}_y\hat{p}_z\hat{L}_x\hat{L}_z + i\hbar\hat{p}_y\hat{p}_z\hat{L}_y \\ &= \hat{p}_z\hat{L}_x\hat{p}_y\hat{L}_z + i\hbar\hat{p}_y\hat{p}_z\hat{L}_y - i\hbar\hat{p}_z\hat{p}_z\hat{L}_z; \end{aligned} \tag{I.93}$$

$$\begin{aligned} \hat{p}_z\hat{L}_x\hat{p}_z\hat{L}_y &= \hat{p}_z\hat{p}_z\hat{L}_x\hat{L}_y - i\hbar\hat{p}_z\hat{p}_y\hat{L}_y \\ &= \hat{p}_z\hat{p}_z\hat{L}_y\hat{L}_x - i\hbar\hat{p}_z\hat{p}_y\hat{L}_y + i\hbar\hat{p}_z\hat{p}_z\hat{L}_z \\ &= \hat{p}_z\hat{L}_y\hat{p}_z\hat{L}_x - i\hbar\hat{p}_z\hat{p}_y\hat{L}_y + i\hbar\hat{p}_z\hat{p}_z\hat{L}_z - i\hbar\hat{p}_z\hat{p}_x\hat{L}_x; \end{aligned} \tag{I.94}$$

$$\begin{aligned} \hat{p}_x\hat{L}_z\hat{p}_y\hat{L}_z &= \hat{p}_x\hat{p}_y\hat{L}_z\hat{L}_z - i\hbar\hat{p}_x\hat{p}_x\hat{L}_z \\ &= \hat{p}_y\hat{L}_z\hat{p}_x\hat{L}_z - i\hbar\hat{p}_x\hat{p}_x\hat{L}_z - i\hbar\hat{p}_y\hat{p}_y\hat{L}_z; \end{aligned} \tag{I.95}$$

$$\begin{aligned} \hat{p}_z\hat{L}_y\hat{p}_x\hat{L}_z &= \hat{p}_z\hat{p}_x\hat{L}_y\hat{L}_z - i\hbar\hat{p}_z\hat{p}_z\hat{L}_z \\ &= \hat{p}_x\hat{L}_z\hat{p}_z\hat{L}_y - i\hbar\hat{p}_z\hat{p}_z\hat{L}_z + i\hbar\hat{p}_x\hat{p}_x\hat{L}_x. \end{aligned} \tag{I.96}$$

Substitution in Eq. (I.92) yields, after canceling,

$$\begin{aligned} &i\hbar\Big(\hat{p}_y\hat{p}_z\hat{L}_y - \hat{p}_z^2\hat{L}_z - \hat{p}_z\hat{p}_y\hat{L}_y + \hat{p}_z^2\hat{L}_z - \hat{p}_z\hat{p}_x\hat{L}_x \\ &\quad - \hat{p}_x^2\hat{L}_z - \hat{p}_y^2\hat{L}_z - \hat{p}_z^2\hat{L}_z + \hat{p}_z\hat{p}_x\hat{L}_x\Big) \\ &= -i\hbar\left(\hat{p}_x^2\hat{L}_z + \hat{p}_y^2\hat{L}_z + \hat{p}_z^2\hat{L}_z\right) = -i\hbar\hat{\mathbf{p}}^2\hat{L}_z, \end{aligned} \tag{I.97}$$

for Eq. (I.82) and Eq. (I.83). In a following step, we perform the multiplications in Eq. (I.84) and Eq. (I.85):

$$\begin{aligned} mk\Big[&-\hat{p}_y\hat{L}_z\frac{y}{r} + \hat{p}_z\hat{L}_y\frac{y}{r} + i\hbar\hat{p}_x\frac{y}{r} - \frac{x}{r}\hat{p}_z\hat{L}_x + \frac{x}{r}\hat{p}_x\hat{L}_z + i\hbar\frac{x}{r}\hat{p}_y \\ &+ \hat{p}_z\hat{L}_x\frac{x}{r} - \hat{p}_x\hat{L}_z\frac{x}{r} - i\hbar\hat{p}_y\frac{x}{r} + \frac{y}{r}\hat{p}_y\hat{L}_z - \frac{y}{r}\hat{p}_z\hat{L}_y - i\hbar\frac{y}{r}\hat{p}_x\Big]. \end{aligned} \tag{I.98}$$

Collecting terms, we obtain

$$mk\Big[-\hat{p}_y\hat{L}_z\frac{y}{r}+\hat{p}_z\hat{L}_y\frac{y}{r}-\frac{x}{r}\hat{p}_z\hat{L}_x+\frac{x}{r}\hat{p}_x\hat{L}_z+\hat{p}_z\hat{L}_x\frac{x}{r}-\hat{p}_x\hat{L}_z\frac{x}{r} \\ +\frac{y}{r}\hat{p}_y\hat{L}_z-\frac{y}{r}\hat{p}_z\hat{L}_y+i\hbar\left(\hat{p}_x\frac{y}{r}+\frac{x}{r}\hat{p}_y-\hat{p}_y\frac{x}{r}-\frac{y}{r}\hat{p}_x\right)\Big]. \tag{I.99}$$

Following the commutation relations $\left[\hat{p}_i,\frac{\hat{r}_j}{r}\right]=-i\hbar\left(\delta_{ij}\frac{1}{r}-\frac{\hat{r}_i\hat{r}_j}{r^3}\right)$, as derived in §I.3.1, we can rewrite the last term as

$$\begin{aligned}\hat{p}_x\frac{y}{r}+\frac{x}{r}\hat{p}_y-\hat{p}_y\frac{x}{r}-\frac{y}{r}\hat{p}_x&=\left(\hat{p}_x\frac{y}{r}-\frac{y}{r}\hat{p}_x\right)-\left(\hat{p}_y\frac{x}{r}-\frac{x}{r}\hat{p}_y\right)\\&=\left[\hat{p}_x,\frac{y}{r}\right]-\left[\hat{p}_y,\frac{x}{r}\right]\\&=i\hbar\frac{xy}{r^3}-i\hbar\frac{yx}{r^3}=0.\end{aligned} \tag{I.100}$$

The expression in Eq. (I.99) can be simplified further using the commutation relations $\left[\hat{p}_i,\hat{r}_j/r\right]=-i\hbar\left(\delta_{ij}1/r-\hat{r}_i\hat{r}_j/r^3\right)$ and $\left[\hat{L}_i,\hat{r}_j/r\right]=i\hbar\varepsilon_{ijk}\hat{r}_k/r$. Rewriting the first four terms as

$$\begin{aligned}-\hat{p}_y\hat{L}_z\frac{y}{r}&=-\hat{p}_y\frac{y}{r}\hat{L}_z+i\hbar\hat{p}_y\frac{x}{r}\\&=-\frac{y}{r}\hat{p}_y\hat{L}_z+i\hbar\left(\frac{1}{r}-\frac{y^2}{r^3}\right)\hat{L}_z+i\hbar\hat{p}_y\frac{x}{r},\end{aligned} \tag{I.101}$$

$$\hat{p}_z\hat{L}_y\frac{y}{r}=\hat{p}_z\frac{y}{r}\hat{L}_y=\frac{y}{r}\hat{p}_z\hat{L}_y+i\hbar\frac{yz}{r^3}\hat{L}_y, \tag{I.102}$$

$$-\frac{x}{r}\hat{p}_z\hat{L}_x=-\hat{p}_z\frac{x}{r}\hat{L}_x+i\hbar\frac{xz}{r^3}\hat{L}_x=-\hat{p}_z\hat{L}_x\frac{x}{r}+i\hbar\frac{xz}{r^3}\hat{L}_x, \tag{I.103}$$

$$\frac{x}{r}\hat{p}_x\hat{L}_z=\hat{p}_x\frac{x}{r}\hat{L}_z+i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{L}_z$$

and

$$=\hat{p}_x\hat{L}_z\frac{x}{r}-i\hbar\hat{p}_x\frac{y}{r}+i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{L}_z, \tag{I.104}$$

yields for Eq. (I.99)

$$mk\Big[-\frac{y}{r}\hat{p}_y\hat{L}_z+i\hbar\left(\frac{1}{r}-\frac{y^2}{r^3}\right)\hat{L}_z+i\hbar\hat{p}_y\frac{x}{r}+\frac{y}{r}\hat{p}_z\hat{L}_y+i\hbar\frac{yz}{r^3}\hat{L}_y-\hat{p}_z\hat{L}_x\frac{x}{r} \\ +i\hbar\frac{xz}{r^3}\hat{L}_x+\hat{p}_x\hat{L}_z\frac{x}{r}-i\hbar\hat{p}_x\frac{y}{r}+i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{L}_z+\hat{p}_z\hat{L}_x\frac{x}{r}-\hat{p}_x\hat{L}_z\frac{x}{r} \\ +\frac{y}{r}\hat{p}_y\hat{L}_z-\frac{y}{r}\hat{p}_z\hat{L}_y\Big]. \tag{I.105}$$

On canceling, this reduces to

$$i\hbar mk\left[\hat{p}_y\frac{x}{r}-\hat{p}_x\frac{y}{r}+\frac{yz}{r^3}\hat{L}_y+\frac{xz}{r^3}\hat{L}_x+\left(\frac{1}{r}-\frac{y^2}{r^3}\right)\hat{L}_z+\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{L}_z\right]. \tag{I.106}$$

We can rewrite the first two terms as follows:

$$\begin{aligned}\hat{p}_y\frac{x}{r}-\hat{p}_x\frac{y}{r}&=\frac{x}{r}\hat{p}_y+i\hbar\frac{xy}{r^3}-\frac{y}{r}\hat{p}_x-i\hbar\frac{yx}{r^3}\\&=\frac{1}{r}\left(x\hat{p}_y-y\hat{p}_x\right)=\frac{1}{r}\hat{L}_z.\end{aligned} \tag{I.107}$$

The third and fourth terms equal

$$\begin{aligned}\frac{yz}{r^3}\hat{L}_y+\frac{xz}{r^3}\hat{L}_x&=\frac{z}{r^3}\left[y\hat{L}_y+x\hat{L}_x\right]\\&=\frac{z}{r^3}\left[y\left(z\hat{p}_x-x\hat{p}_z\right)+x\left(y\hat{p}_z-z\hat{p}_y\right)\right]\\&=\frac{z}{r^3}\left[yz\hat{p}_x-xy\hat{p}_z+xy\hat{p}_z-xz\hat{p}_y\right]\\&=-\frac{z}{r^3}\left[z\left(x\hat{p}_y-y\hat{p}_x\right)\right]\\&=-\frac{z^2}{r^3}\hat{L}_z.\end{aligned}\tag{I.108}$$

Substitution of these results in Eq. (I.106) yields

$$\begin{aligned}&i\hbar mk\left[\frac{1}{r}\hat{L}_z-\frac{z^2}{r^3}\hat{L}_z+\left(\frac{1}{r}-\frac{y^2}{r^3}\right)\hat{L}_z+\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{L}_z\right]\\&=i\hbar mk\left[\frac{3}{r}\hat{L}_z-\frac{1}{r^3}\left(x^2+y^2+z^2\right)\hat{L}_z\right]\\&=i\hbar mk\left[\frac{3}{r}\hat{L}_z-\frac{r^2}{r^3}\hat{L}_z\right]=i\hbar mk\frac{2}{r}\hat{L}_z.\end{aligned}\tag{I.109}$$

After inserting the results from Eq. (I.97) and Eq. (I.109) into Eqs. (I.82) through (I.86), we end up with the following commutation relation:

$$\begin{aligned}\left[\hat{M}_x,\hat{M}_y\right]&=-i\hbar\hat{\mathbf{p}}^2\hat{L}_z+i\hbar mk\frac{2}{r}\hat{L}_z\\&=i\hbar(-2m)\left(\frac{\hat{\mathbf{p}}^2}{2m}-\frac{k}{r}\right)\hat{L}_z\\&=i\hbar(-2m\mathscr{H})\hat{L}_z,\end{aligned}\tag{I.110}$$

where we have introduced the Hamiltonian $\mathscr{H}$ in the last line. Naturally, $\left[\hat{M}_x,\hat{M}_x\right]=0$. The remaining commutation relations can be determined by cyclic permutation of x, y, and z, and they result in the generalized commutation relation

$$\left[\hat{M}_i,\hat{M}_j\right]=i\hbar(-2m\mathscr{H})\,\varepsilon_{ijk}\hat{L}_k.\tag{I.111}$$

I.6 MORE AUXILIARY COMMUTATORS

In our evaluation of the commutators $\left[\hat{L}_i,\hat{\mathscr{H}}\right]$ and $\left[\hat{M}_i,\hat{\mathscr{H}}\right]$ (see §I.7), recourse will be needed to two more auxiliary commutators, which we need to derive first.

I.6.1 Commutation relations for $\left[\hat{p}_i,\hat{\mathscr{H}}\right]$

The commutation relations for $\left[\hat{p}_i,\hat{\mathscr{H}}\right]$ are easily derived. The Hamiltonian for the Coulomb problem was given in Eq. (9.37) and can be used to write

$$\left[\hat{p}_i,\hat{\mathscr{H}}\right]=\left[\hat{p}_i,\frac{\hat{p}^2}{2m}-\frac{k}{r}\right]=\frac{1}{2m}\left[\hat{p}_i,\sum_{j=1}^{3}\hat{p}_j^2\right]-k\left[\hat{p}_i,\frac{1}{r}\right].\tag{I.112}$$

Given that the momentum operators commute with themselves (i.e., $\left[\hat{p}_i,\hat{p}_j\right]=0$), the first commutator in Eq. (I.112) vanishes. The second commutator yields

$$\left[\hat{p}_i,\frac{1}{r}\right]=\hat{p}_i\frac{1}{r}-\frac{1}{r}\hat{p}_i=-i\hbar\frac{\partial}{\partial\hat{r}_i}\frac{1}{r}-i\hbar\frac{1}{r}\frac{\partial}{\partial\hat{r}_i}+i\hbar\frac{1}{r}\frac{\partial}{\partial\hat{r}_i}=-i\hbar\frac{\partial}{\partial\hat{r}_i}\frac{1}{r},\tag{I.113}$$

where $\partial/\partial\hat{r}_i(1/r)$ can be written as

$$\frac{\partial}{\partial \hat{r}_i}\frac{1}{r} = \frac{\partial\left(\sum_{j=1}^{3}\hat{r}_j^2\right)^{-\frac{1}{2}}}{\partial\sum_{j=1}^{3}\hat{r}_j^2}\frac{\partial\sum_{j=1}^{3}\hat{r}_j^2}{\partial\hat{r}_i} = -\frac{1}{2}\left(\sum_{j=1}^{3}\hat{r}_j^2\right)^{-\frac{3}{2}} 2\hat{r}_i = -\frac{\hat{r}_i}{r^3}. \tag{I.114}$$

Substitution of these results in Eq. (I.112) yields the following commutation relations for $\left[\hat{p}_i, \hat{\mathscr{H}}\right]$:

$$\left[\hat{p}_i, \hat{\mathscr{H}}\right] = -i\hbar k\frac{\hat{r}_i}{r^3}. \tag{I.115}$$

I.6.2 Commutation relations for $\left[\hat{p}_i^2, \hat{r}_j/r\right]$

From the commutator identity $\left[\hat{A}\hat{B},\hat{C}\right] = \left[\hat{A},\hat{C}\right]\hat{B} + \hat{A}\left[\hat{B},\hat{C}\right]$, we can write

$$\left[\hat{p}_i^2, \frac{\hat{r}_j}{r}\right] = \left[\hat{p}_i, \frac{\hat{r}_j}{r}\right]\hat{p}_i + \hat{p}_i\left[\hat{p}_i, \frac{\hat{r}_j}{r}\right]. \tag{I.116}$$

Following Eq. (I.41), we obtain

$$\left[\hat{p}_i^2, \frac{\hat{r}_j}{r}\right] = -i\hbar\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\hat{p}_i - i\hbar\hat{p}_i\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right). \tag{I.117}$$

Acting with the second term on a test function f yields

$$\begin{aligned}-i\hbar\hat{p}_i\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\cdot f &= -i\hbar\left(-i\hbar\frac{\partial}{\partial\hat{r}_i}\right)\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\cdot f \\ &= -\hbar^2\left[\frac{\partial}{\partial\hat{r}_i}\left(\delta_{ij}\frac{1}{r}\cdot f\right) - \frac{\partial}{\partial\hat{r}_i}\left(\frac{\hat{r}_i\hat{r}_j}{r^3}\cdot f\right)\right].\end{aligned} \tag{I.118}$$

The first partial derivative in Eq. (I.118) can be written as

$$\begin{aligned}\frac{\partial}{\partial\hat{r}_i}\left(\delta_{ij}\frac{1}{r}\cdot f\right) &= \delta_{ij}\frac{1}{r}\frac{\partial f}{\partial\hat{r}_i} + \delta_{ij}f\frac{\partial}{\partial\hat{r}_i}\frac{1}{r} \\ &= \delta_{ij}\frac{1}{r}\frac{\partial f}{\partial\hat{r}_i} - \delta_{ij}\frac{\hat{r}_i}{r^3}f,\end{aligned} \tag{I.119}$$

where use was made of Eq. (I.114). The second derivative in Eq. (I.118) gives

$$\begin{aligned}\frac{\partial}{\partial\hat{r}_i}\left(\frac{\hat{r}_i\hat{r}_j}{r^3}\cdot f\right) &= \frac{1}{r^3}\frac{\partial}{\partial\hat{r}_i}\left(\hat{r}_i\hat{r}_j\cdot f\right) + \hat{r}_i\hat{r}_j f\frac{\partial}{\partial\hat{r}_i}\frac{1}{r^3} \\ &= \frac{\hat{r}_i\hat{r}_j}{r^3}\frac{\partial f}{\partial\hat{r}_i} + \frac{1}{r^3}f\frac{\partial}{\partial\hat{r}_i}\left(\hat{r}_i\hat{r}_j\right) - \hat{r}_i\hat{r}_j f\frac{3}{2}r^{-5}2\hat{r}_i \\ &= \frac{\hat{r}_i\hat{r}_j}{r^3}\frac{\partial f}{\partial\hat{r}_i} + \frac{\hat{r}_i}{r^3}f\frac{\partial\hat{r}_j}{\partial\hat{r}_i} + \frac{\hat{r}_j}{r^3}f\frac{\partial\hat{r}_i}{\partial\hat{r}_i} - 3\frac{\hat{r}_i^2\hat{r}_j}{r^5}f \\ &= \frac{\hat{r}_i\hat{r}_j}{r^3}\frac{\partial f}{\partial\hat{r}_i} + \delta_{ij}\frac{\hat{r}_i}{r^3}f + \frac{\hat{r}_j}{r^3}f - 3\frac{\hat{r}_i^2\hat{r}_j}{r^5}f.\end{aligned} \tag{I.120}$$

Substitution of these results in Eq. (I.118), and removal of f, yields

$$-i\hbar\hat{p}_i\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right) = -\hbar^2\left[\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\frac{\partial}{\partial\hat{r}_i} - 2\delta_{ij}\frac{\hat{r}_i}{r^3} - \frac{\hat{r}_j}{r^3} + 3\frac{\hat{r}_i^2\hat{r}_j}{r^5}\right], \tag{I.121}$$

where

$$\begin{aligned}-\hbar^2\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\frac{\partial}{\partial\hat{r}_i} &= -i\hbar\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\left(-i\hbar\frac{\partial}{\partial\hat{r}_i}\right) \\ &= -i\hbar\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\hat{p}_i.\end{aligned} \tag{I.122}$$

Substitution of Eq. (I.121) and Eq. (I.122) into Eq. (I.117) finally yields the commutation relations for $\left[\hat{p}_i^2, \frac{\hat{r}_j}{r}\right]$:

$$\left[\hat{p}_i^2, \frac{\hat{r}_j}{r}\right] = -2i\hbar\left(\delta_{ij}\frac{1}{r} - \frac{\hat{r}_i\hat{r}_j}{r^3}\right)\hat{p}_i + \hbar^2\left(2\delta_{ij}\frac{\hat{r}_i}{r^3} + \frac{\hat{r}_j}{r^3} - 3\frac{\hat{r}_i^2\hat{r}_j}{r^5}\right). \tag{I.123}$$

When $i = j$, this becomes

$$\left[\hat{p}_i^2, \frac{\hat{r}_i}{r}\right] = -2i\hbar\left(\frac{1}{r} - \frac{\hat{r}_i^2}{r^3}\right)\hat{p}_i + \hbar^2\left(\frac{3\hat{r}_i}{r^3} - \frac{3\hat{r}_i^3}{r^5}\right). \tag{I.124}$$

For $i \neq j$, on the other hand, we have

$$\left[\hat{p}_i^2, \frac{\hat{r}_j}{r}\right] = 2i\hbar\frac{\hat{r}_i\hat{r}_j}{r^3}\hat{p}_i + \hbar^2\left(\frac{\hat{r}_j}{r^3} - 3\frac{\hat{r}_i^2\hat{r}_j}{r^5}\right). \tag{I.125}$$

I.7 COMMUTATORS WITH THE HAMILTONIAN

I.7.1 Commutation relations for $\left[\hat{L}_i, \hat{\mathscr{H}}\right]$

We want to prove that the components $\hat{L}_i$ of the angular momentum operator $\hat{L}$ commute with the Hamiltonian $\hat{\mathscr{H}}$ for the Kepler problem:

$$\left[\hat{L}_i, \hat{\mathscr{H}}\right] = 0, \qquad \forall i = 1 \to 3. \tag{I.126}$$

From Eq. (9.37), we can write

$$\left[\hat{L}_i, \hat{\mathscr{H}}\right] = \left[\hat{L}_i, \frac{\hat{p}^2}{2m} - \frac{k}{r}\right] = \frac{1}{2m}\left[\hat{L}_i, \sum_{j=1}^{3}\hat{p}_j^2\right] - k\left[\hat{L}_i, \frac{1}{r}\right]. \tag{I.127}$$

Let us therefore verify the validity of Eq. (I.126) by investigating the commutation relations for the components of $\hat{L}$ with the kinetic energy operator $\hat{T}$ and the potential energy operator $\hat{V}$ listed previously.

The commutator with the kinetic energy operator is easily obtained. As an example, for L_x, we have

$$\begin{aligned}\left[\hat{L}_x, \hat{p}^2\right] &= \left[\hat{L}_x, \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2\right]\\ &= \left[\hat{L}_x, \hat{p}_x^2\right] + \left[\hat{L}_x, \hat{p}_y^2\right] + \left[\hat{L}_x, \hat{p}_z^2\right].\end{aligned} \tag{I.128}$$

From the commutator identity $\left[\hat{A}, \hat{B}\hat{C}\right] = \left[\hat{A}, \hat{C}\right]\hat{B} + \hat{C}\left[\hat{A}, \hat{B}\right]$, and the commutation relations of Eq. (I.20), we obtain

$$\begin{aligned}\left[\hat{L}_x, \hat{p}^2\right] &= \left[\hat{L}_x, \hat{p}_y\right]\hat{p}_y + \hat{p}_y\left[\hat{L}_x, \hat{p}_y\right] + \left[\hat{L}_x, \hat{p}_z\right]\hat{p}_z + \hat{p}_z\left[\hat{L}_x, \hat{p}_z\right]\\ &= i\hbar\hat{p}_z\hat{p}_y + i\hbar\hat{p}_y\hat{p}_z - i\hbar\hat{p}_y\hat{p}_z - i\hbar\hat{p}_z\hat{p}_y = 0.\end{aligned} \tag{I.129}$$

Similarly, we can prove that $\hat{L}_x$ commutes with the potential energy operator:

$$\begin{aligned}\left[\hat{L}_x, \frac{1}{r}\right] &= \left[y\hat{p}_z - z\hat{p}_y, \frac{1}{r}\right] = y\left[\hat{p}_z, \frac{1}{r}\right] - z\left[\hat{p}_y, \frac{1}{r}\right]\\ &= i\hbar\frac{yz}{r^3} - i\hbar\frac{zy}{r^3} = 0,\end{aligned} \tag{I.130}$$

where we have used Eq. (I.113) and Eq. (I.114) in the second line. Substitution of Eq. (I.129) and Eq. (I.130) into Eq. (I.127) shows that $\hat{L}_x$ indeed commutes with the Hamiltonian. This can be generalized via cyclic permutation of x, y, and z to yield Eq. (I.126).

I.7.2 Commutation relations for $\left[\hat{M}_i, \hat{\mathscr{H}}\right]$

To derive the commutation relations between the LRL vector $\hat{\mathbf{M}}$ and the Hamiltonian $\hat{\mathscr{H}}$, it will be convenient to use the alternative form of $\hat{\mathbf{M}}$, as given in Eq. (I.64). Our aim is to demonstrate the *commutativity* of the components of $\hat{\mathbf{M}}$ with the Hamiltonian $\hat{\mathscr{H}}$ for the Kepler problem:

$$\left[\hat{M}_i, \hat{\mathscr{H}}\right] = 0, \qquad \forall i = 1 \to 3. \tag{I.131}$$

Let us, by way of example, evaluate the commutator of the Hamiltonian with the x-component of the LRL vector:

$$\begin{aligned}\left[\hat{M}_x, \hat{\mathscr{H}}\right] &= \left[\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)_x - mk\frac{x}{r}, \hat{\mathscr{H}}\right] \\ &= \left[\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}}\right)_x, \hat{\mathscr{H}}\right] - i\hbar\left[\hat{p}_x, \hat{\mathscr{H}}\right] - mk\left[\frac{x}{r}, \hat{\mathscr{H}}\right].\end{aligned} \tag{I.132}$$

Let us compute each of the three terms individually. The first commutator in Eq. (I.132) can be written as

$$\begin{aligned}\left[\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}}\right)_x, \hat{\mathscr{H}}\right] &= \left[\hat{p}_y\hat{L}_z - \hat{p}_z\hat{L}_y, \hat{\mathscr{H}}\right] \\ &= \left[\hat{p}_y\hat{L}_z, \hat{\mathscr{H}}\right] - \left[\hat{p}_z\hat{L}_y, \hat{\mathscr{H}}\right] \\ &= \left[\hat{p}_y, \hat{\mathscr{H}}\right]\hat{L}_z + \hat{p}_y\left[\hat{L}_z, \hat{\mathscr{H}}\right] - \left[\hat{p}_z, \hat{\mathscr{H}}\right]\hat{L}_y - \hat{p}_z\left[\hat{L}_y, \hat{\mathscr{H}}\right],\end{aligned} \tag{I.133}$$

where we have used the commutator identity $\left[\hat{A}\hat{B}, \hat{C}\right] = \left[\hat{A}, \hat{C}\right]\hat{B} + \hat{A}\left[\hat{B}, \hat{C}\right]$ in the last line. We also know from the previous section that the Hamiltonian commutes with the components of the angular momentum—that is, $\left[\hat{L}_z, \hat{\mathscr{H}}\right] = \left[\hat{L}_y, \hat{\mathscr{H}}\right] = 0$. This simplifies Eq. (I.133) to

$$\left[\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}}\right)_x, \hat{\mathscr{H}}\right] = \left[\hat{p}_y, \hat{\mathscr{H}}\right]\hat{L}_z - \left[\hat{p}_z, \hat{\mathscr{H}}\right]\hat{L}_y. \tag{I.134}$$

The remaining two commutators were determined in §I.6. From Eq. (I.115), we thus obtain

$$\begin{aligned}\left[\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}}\right)_x, \hat{\mathscr{H}}\right] &= -i\hbar k\frac{y}{r^3}\hat{L}_z + i\hbar k\frac{z}{r^3}\hat{L}_y \\ &= -i\hbar k\left[\frac{y}{r^3}\left(x\hat{p}_y - y\hat{p}_x\right) - \frac{z}{r^3}\left(z\hat{p}_x - x\hat{p}_z\right)\right] \\ &= -i\hbar k\left[\frac{xy}{r^3}\hat{p}_y - \frac{y^2}{r^3}\hat{p}_x - \frac{z^2}{r^3}\hat{p}_x + \frac{xz}{r^3}\hat{p}_z\right].\end{aligned} \tag{I.135}$$

Addition and subtraction of $\frac{x^2}{r^3}\hat{p}_x$ inside the brackets finally yields

$$\begin{aligned}\left[\left(\hat{\mathbf{p}} \times \hat{\mathbf{L}}\right)_x, \hat{\mathscr{H}}\right] &= -i\hbar k\left[\frac{xy}{r^3}\hat{p}_y + \frac{xz}{r^3}\hat{p}_z + \frac{x^2}{r^3}\hat{p}_x - \left(\frac{y^2}{r^3}\hat{p}_x + \frac{z^2}{r^3}\hat{p}_x + \frac{x^2}{r^3}\hat{p}_x\right)\right] \\ &= -i\hbar k\left[\frac{xy}{r^3}\hat{p}_y + \frac{xz}{r^3}\hat{p}_z + \frac{x^2}{r^3}\hat{p}_x - \frac{1}{r}\hat{p}_x\right].\end{aligned} \tag{I.136}$$

The second commutator in Eq. (I.132) equals, by Eq. (I.115),

$$\left[\hat{p}_x, \hat{\mathscr{H}}\right] = -i\hbar k\frac{x}{r^3}. \tag{I.137}$$

The third and final commutator in Eq. (I.132) can be evaluated as follows:

$$\begin{aligned}\left[\frac{x}{r}, \hat{\mathscr{H}}\right] = \left[\frac{x}{r}, \frac{\hat{p}^2}{2m} - \frac{k}{r}\right] &= \left[\frac{x}{r}, \frac{1}{2m}\left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2\right)\right] - \left[\frac{x}{r}, \frac{k}{r}\right] \\ &= \frac{1}{2m}\left\{\left[\frac{x}{r}, \hat{p}_x^2\right] + \left[\frac{x}{r}, \hat{p}_y^2\right] + \left[\frac{x}{r}, \hat{p}_z^2\right]\right\}.\end{aligned} \tag{I.138}$$

The three commutators in Eq. (I.138) were derived in §I.6. From Eq. (I.124) and Eq. (I.125), and the fact that $\left[\hat{A},\hat{B}\right]=-\left[\hat{B},\hat{A}\right]$, we obtain

$$\begin{aligned}\left[\frac{x}{r},\hat{\mathscr{H}}\right]=\frac{1}{2m}\Bigg\{&2i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{p}_x-\hbar^2\left(\frac{3x}{r^3}-\frac{3x^3}{r^5}\right)\\&-2i\hbar\frac{xy}{r^3}\hat{p}_y-\hbar^2\left(\frac{x}{r^3}-3\frac{xy^2}{r^5}\right)\\&-2i\hbar\frac{xz}{r^3}\hat{p}_z-\hbar^2\left(\frac{x}{r^3}-3\frac{xz^2}{r^5}\right)\Bigg\}.\end{aligned}\tag{I.139}$$

This can be written as

$$\begin{aligned}\left[\frac{x}{r},\hat{\mathscr{H}}\right]=\frac{1}{2m}\Bigg\{&2i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{p}_x-\hbar^2\left[\frac{5x}{r^3}-\frac{3x}{r^5}\left(x^2+y^2+z^2\right)\right]\\&-2i\hbar\frac{xy}{r^3}\hat{p}_y-2i\hbar\frac{xz}{r^3}\hat{p}_z\Bigg\},\end{aligned}\tag{I.140}$$

and since $r^2=x^2+y^2+z^2$, we obtain

$$\begin{aligned}\left[\frac{x}{r},\hat{\mathscr{H}}\right]&=\frac{1}{2m}\left\{2i\hbar\left(\frac{1}{r}-\frac{x^2}{r^3}\right)\hat{p}_x-2\hbar^2\frac{x}{r^3}-2i\hbar\frac{xy}{r^3}\hat{p}_y-2i\hbar\frac{xz}{r^3}\hat{p}_z\right\}\\&=\frac{i\hbar}{m}\left\{\frac{1}{r}\hat{p}_x-\frac{x^2}{r^3}\hat{p}_x-\frac{xy}{r^3}\hat{p}_y-\frac{xz}{r^3}\hat{p}_z+i\hbar\frac{x}{r^3}\right\}.\end{aligned}\tag{I.141}$$

Substitution of Eq. (I.136), Eq. (I.137), and Eq. (I.141) into Eq. (I.132) finally yields

$$\begin{aligned}\left[\hat{M}_x,\hat{\mathscr{H}}\right]=&-i\hbar k\left[\frac{xy}{r^3}\hat{p}_y+\frac{xz}{r^3}\hat{p}_z+\frac{x^2}{r^3}\hat{p}_x-\frac{1}{r}\hat{p}_x\right]-\hbar^2k\frac{x}{r^3}\\&-i\hbar k\left[\frac{1}{r}\hat{p}_x-\frac{x^2}{r^3}\hat{p}_x-\frac{xy}{r^3}\hat{p}_y-\frac{xz}{r^3}\hat{p}_z+i\hbar\frac{x}{r^3}\right]=0.\end{aligned}\tag{I.142}$$

The remaining commutation relations can be determined by cyclic permutation of x, y, and z, and they result in the generalized commutation relations

$$\left[\hat{M}_i,\hat{\mathscr{H}}\right]=0,\qquad\forall i=1\rightarrow 3.\tag{I.143}$$

This validates our assertion of the commutativity of $\hat{\mathbf{M}}$ and $\hat{\mathscr{H}}$.

I.8 ANGULAR MOMENTUM COMMUTATION RELATIONS

Given the defining relation $\hat{L}_{ij}=\hat{r}_i\hat{p}_j-\hat{r}_j\hat{p}_i$, we can write

$$\left[\hat{L}_{ij},\hat{L}_{kl}\right]=\left[\left(\hat{r}_i\hat{p}_j-\hat{r}_j\hat{p}_i\right),\left(\hat{r}_k\hat{p}_l-\hat{r}_l\hat{p}_k\right)\right].\tag{I.144}$$

Expanding yields

$$\left[\hat{L}_{ij},\hat{L}_{kl}\right]=\left[\hat{r}_i\hat{p}_j,\hat{r}_k\hat{p}_l\right]-\left[\hat{r}_j\hat{p}_i,\hat{r}_k\hat{p}_l\right]-\left[\hat{r}_i\hat{p}_j,\hat{r}_l\hat{p}_k\right]+\left[\hat{r}_j\hat{p}_i,\hat{r}_l\hat{p}_k\right].\tag{I.145}$$

The four terms in Eq. (I.145) can be expanded further using the commutator identities $\left[\hat{A},\hat{B}\hat{C}\right]=\left[\hat{A},\hat{B}\right]\hat{C}+\hat{B}\left[\hat{A},\hat{C}\right]$ and $\left[\hat{A}\hat{B},\hat{C}\right]=\left[\hat{A},\hat{C}\right]\hat{B}+\hat{A}\left[\hat{B},\hat{C}\right]$, as well as the commutation relations $\left[\hat{r}_i,\hat{r}_j\right]=\left[\hat{p}_i,\hat{p}_j\right]=0$ (see §I.1):

$$\begin{aligned}\left[\hat{r}_i\hat{p}_j,\hat{r}_k\hat{p}_l\right]&=\left[\hat{r}_i\hat{p}_j,\hat{r}_k\right]\hat{p}_l+\hat{r}_k\left[\hat{r}_i\hat{p}_j,\hat{p}_l\right]\\&=\left[\hat{r}_i,\hat{r}_k\right]\hat{p}_j\hat{p}_l+\hat{r}_i\left[\hat{p}_j,\hat{r}_k\right]\hat{p}_l+\hat{r}_k\left[\hat{r}_i,\hat{p}_l\right]\hat{p}_j+\hat{r}_k\hat{r}_i\left[\hat{p}_j,\hat{p}_l\right]\\&=-\hat{r}_i\left[\hat{r}_k,\hat{p}_j\right]\hat{p}_l+\hat{r}_k\left[\hat{r}_i,\hat{p}_l\right]\hat{p}_j.\end{aligned}\tag{I.146}$$

Given Heisenberg's canonical commutation relations $\left[\hat{r}_i, \hat{p}_j\right] = i\hbar\delta_{ij}$, this becomes

$$\left[\hat{r}_i\hat{p}_j, \hat{r}_k\hat{p}_l\right] = -i\hbar\delta_{jk}\hat{r}_i\hat{p}_l + i\hbar\delta_{il}\hat{r}_k\hat{p}_j. \tag{I.147}$$

Substitution in Eq. (I.145) finally yields

$$\begin{aligned}\left[\hat{L}_{ij}, \hat{L}_{kl}\right] = i\hbar\big[&\delta_{jk}\left(\hat{r}_l\hat{p}_i - \hat{r}_i\hat{p}_l\right) + \delta_{il}\left(\hat{r}_k\hat{p}_j - \hat{r}_j\hat{p}_k\right) \\ &+ \delta_{ik}\left(\hat{r}_j\hat{p}_l - \hat{r}_l\hat{p}_j\right) + \delta_{jl}\left(\hat{r}_i\hat{p}_k - \hat{r}_k\hat{p}_i\right)\big].\end{aligned} \tag{I.148}$$

Clearly, if $i \neq j \neq k \neq l$, then

$$\left[\hat{L}_{ij}, \hat{L}_{kl}\right] = 0. \tag{I.149}$$

If, instead, $i = k$, we have

$$\left[\hat{L}_{ij}, \hat{L}_{il}\right] = i\hbar\left(\hat{r}_j\hat{p}_l - \hat{r}_l\hat{p}_j\right) = i\hbar\hat{L}_{jl}. \tag{I.150}$$

Changing the dummy variable l into k, we obtain

$$\left[\hat{L}_{ij}, \hat{L}_{ik}\right] = i\hbar\hat{L}_{jk}. \tag{I.151}$$

Summarizing:

$$\left[\hat{L}_{ij}, \hat{L}_{ik}\right] = i\hbar\hat{L}_{jk}, \qquad \left[\hat{L}_{ij}, \hat{L}_{kl}\right] = 0 \quad \text{if } i \neq j \neq k \neq l. \tag{I.152}$$

Appendix J

Identities of the Laplace-Runge-Lenz vector

J.1 THE SCALAR PRODUCT $\hat{\mathbf{L}}\cdot\hat{\mathbf{M}}$

Here, we want to prove the quantum mechanical relation

$$\hat{\mathbf{L}}\cdot\hat{\mathbf{M}}=0. \tag{J.1}$$

To this aim, we can write, with the help of Eq. (9.93),

$$\begin{aligned}\hat{\mathbf{L}}\cdot\hat{\mathbf{M}}&=\hat{\mathbf{L}}\cdot\left(\frac{1}{2}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)-mk\frac{\hat{\mathbf{r}}}{r}\right)\\&=\frac{1}{2}\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)-mk\hat{\mathbf{L}}\cdot\frac{\hat{\mathbf{r}}}{r}.\end{aligned} \tag{J.2}$$

The first term can be written as

$$\begin{aligned}\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)&=\sum_{i=1}^{3}\hat{L}_i\left(\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i-\left(\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)_i\right)\\&=\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\hat{L}_i\left(\hat{p}_j\hat{L}_k-\hat{L}_j\hat{p}_k\right),\end{aligned} \tag{J.3}$$

where the tensor notation in Eq. (K.21) for the cross product of $\hat{\mathbf{p}}$ and $\hat{\mathbf{L}}$ has been used twice. This can be further expanded as

$$\begin{aligned}\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)&=\sum_{ijk}\varepsilon_{ijk}\hat{L}_i\left(\hat{p}_j\left(\hat{\mathbf{r}}\times\hat{\mathbf{p}}\right)_k-\left(\hat{\mathbf{r}}\times\hat{\mathbf{p}}\right)_j\hat{p}_k\right)\\&=\sum_{ijklm}\hat{L}_i\varepsilon_{ijk}\left(\varepsilon_{klm}\hat{p}_j\hat{r}_l\hat{p}_m-\varepsilon_{jlm}\hat{r}_l\hat{p}_m\hat{p}_k\right).\end{aligned} \tag{J.4}$$

Eq. (J.4) contains two products of Levi-Civita symbols ($\varepsilon_{ijk}\varepsilon_{klm}$ and $\varepsilon_{ijk}\varepsilon_{jlm}$), which can be written in terms of Kronecker deltas with the help of the contracted epsilon identity from Eq. (K.5):

$$\varepsilon_{ijk}\varepsilon_{klm}=\varepsilon_{kij}\varepsilon_{klm}=\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}; \tag{J.5}$$

$$\varepsilon_{ijk}\varepsilon_{jlm}=\varepsilon_{jki}\varepsilon_{jlm}=\delta_{kl}\delta_{im}-\delta_{km}\delta_{il}. \tag{J.6}$$

Substitution in Eq. (J.4) yields

$$\begin{aligned}\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)=&\sum_{ijlm}\hat{L}_i\left(\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}\right)\hat{p}_j\hat{r}_l\hat{p}_m\\&-\sum_{iklm}\hat{L}_i\left(\delta_{kl}\delta_{im}-\delta_{km}\delta_{il}\right)\hat{r}_l\hat{p}_m\hat{p}_k,\end{aligned} \tag{J.7}$$

which can be written as

$$\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)=\sum_{ij}\hat{L}_i\left(\hat{p}_j\hat{r}_i\hat{p}_j-\hat{p}_j\hat{r}_j\hat{p}_i\right)-\sum_{ik}\hat{L}_i\left(\hat{r}_k\hat{p}_i\hat{p}_k-\hat{r}_i\hat{p}_k\hat{p}_k\right). \tag{J.8}$$

After changing the dummy index k into j in the second term, we obtain

$$\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)=\sum_{ij}\hat{L}_i\left(\hat{p}_j\hat{r}_i\hat{p}_j-\hat{p}_j\hat{r}_j\hat{p}_i-\hat{r}_j\hat{p}_i\hat{p}_j+\hat{r}_i\hat{p}_j\hat{p}_j\right). \tag{J.9}$$

Rewriting $\hat{L}_i$ in tensorial notation finally yields

$$\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)=\sum_{ijkl}\varepsilon_{ikl}\hat{r}_k\hat{p}_l\left(\hat{p}_j\hat{r}_i\hat{p}_j-\hat{p}_j\hat{r}_j\hat{p}_i-\hat{r}_j\hat{p}_i\hat{p}_j+\hat{r}_i\hat{p}_j\hat{p}_j\right). \tag{J.10}$$

Let us now separately evaluate the four terms in the right-hand side of Eq. (J.10). Given the canonical commutation relations for position and momentum $\left[\hat{r}_i,\hat{p}_j\right]=i\hbar\delta_{ij}$, as derived in Eq. (I.8), the four terms can be written as

$$\begin{aligned}\hat{r}_k\hat{p}_l\left(\hat{p}_j\hat{r}_i\right)\hat{p}_j&=\hat{r}_k\left(\hat{p}_l\hat{r}_i\right)\hat{p}_j\hat{p}_j-i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j\\&=\hat{r}_k\hat{r}_i\hat{p}_l\hat{p}_j^2-i\hbar\delta_{il}\hat{r}_k\hat{p}_j^2-i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j;\end{aligned} \tag{J.11}$$

$$\begin{aligned}\hat{r}_k\hat{p}_l\hat{p}_j\left(\hat{r}_j\hat{p}_i\right)&=\hat{r}_k\hat{p}_l\hat{p}_j\hat{p}_i\hat{r}_j+i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j\\&=\hat{r}_k\hat{p}_i\hat{p}_l\hat{p}_j\hat{r}_j+i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j;\end{aligned} \tag{J.12}$$

$$\begin{aligned}\hat{r}_k\hat{p}_l\left(\hat{r}_j\hat{p}_i\right)\hat{p}_j&=\hat{r}_k\hat{p}_l\hat{p}_i\hat{r}_j\hat{p}_j+i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j\\&=\hat{r}_k\hat{p}_i\hat{p}_l\hat{p}_j\hat{r}_j+i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j;\end{aligned} \tag{J.13}$$

$$\hat{r}_k\left(\hat{p}_l\hat{r}_i\right)\hat{p}_j\hat{p}_j=\hat{r}_k\hat{r}_i\hat{p}_l\hat{p}_j^2-i\hbar\delta_{il}\hat{r}_k\hat{p}_j^2. \tag{J.14}$$

Substitution in Eq. (J.10) then yields

$$\begin{aligned}\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)=\sum_{ijkl}\varepsilon_{ikl}\Big(&2\hat{r}_k\hat{r}_i\hat{p}_l\hat{p}_j^2-2\hat{r}_k\hat{p}_i\hat{p}_l\hat{p}_j\hat{r}_j\\&-2i\hbar\delta_{il}\hat{r}_k\hat{p}_j^2-3i\hbar\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j\Big).\end{aligned} \tag{J.15}$$

The first term in Eq. (J.15) vanishes since $\sum_{i,k=1}^{3}\varepsilon_{ikl}\hat{r}_i\hat{r}_k=0$ for all l in view of the fact that $\left[\hat{r}_i,\hat{r}_k\right]=0$. This can be seen most easily by explicit computation; for $l=3$, for instance,

$$\sum_{i,k=1}^{3}\varepsilon_{ik3}\hat{r}_i\hat{r}_k=\hat{r}_1\hat{r}_2-\hat{r}_2\hat{r}_1=0, \tag{J.16}$$

and similarly for $l=1$ and $l=2$. The same applies to the second term of Eq. (J.15), where $\sum_{i,l=1}^{3}\varepsilon_{ikl}\hat{p}_i\hat{p}_l=0$ for all k since $\left[\hat{p}_i,\hat{p}_l\right]=0$. The third term in Eq. (J.15) also vanishes since $\sum_{ijkl}\varepsilon_{ikl}\delta_{il}=\sum_{ijk}\varepsilon_{iki}=0$. We are thus left with the following expression:

$$\begin{aligned}\hat{\mathbf{L}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-\hat{\mathbf{L}}\times\hat{\mathbf{p}}\right)&=-\sum_{ijkl}3i\hbar\varepsilon_{ikl}\delta_{ij}\hat{r}_k\hat{p}_l\hat{p}_j\\&=-\sum_{ikl}3i\hbar\varepsilon_{ikl}\hat{r}_k\hat{p}_l\hat{p}_i=0,\end{aligned} \tag{J.17}$$

since, once again, $\sum_{i,l=1}^{3}\varepsilon_{ikl}\hat{p}_i\hat{p}_l=0$ for all k. The second term in Eq. (J.2) can be written as

$$\hat{\mathbf{L}}\cdot\frac{\hat{\mathbf{r}}}{r}=\sum_{i=1}^{3}\hat{L}_i\frac{\hat{r}_i}{r}=\sum_{i=1}^{3}\left(\hat{\mathbf{r}}\times\hat{\mathbf{p}}\right)_i\frac{\hat{r}_i}{r}=\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\hat{r}_j\hat{p}_k\frac{\hat{r}_i}{r}, \tag{J.18}$$

where we have used the tensor notation of Eq. (K.21) for the cross product of $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$. From the commutation relations $\left[\hat{r}_j, \hat{p}_k\right] = i\hbar\delta_{jk}$, this can be rewritten as

$$\begin{aligned}\hat{\mathbf{L}}\cdot\frac{\hat{\mathbf{r}}}{r} &= \sum_{i,j,k=1}^{3} \varepsilon_{ijk}\hat{p}_k\hat{r}_j\frac{\hat{r}_i}{r} + \sum_{i,j,k=1}^{3} i\hbar\varepsilon_{ijk}\delta_{jk}\frac{r_i}{r}\\ &= \frac{1}{r}\sum_{k=1}^{3}\hat{p}_k\sum_{i,j=1}^{3}\varepsilon_{ijk}\hat{r}_j\hat{r}_i + i\hbar\sum_{i,j=1}^{3}\varepsilon_{ijj}\frac{r_i}{r}.\end{aligned} \tag{J.19}$$

The first term in Eq. (J.19) vanishes since $\sum_{i,j=1}^{3}\varepsilon_{ijk}\hat{r}_j\hat{r}_i = 0$ for all k in view of the fact that $\left[\hat{r}_i, \hat{r}_j\right] = 0$. The second term in Eq. (J.19) also vanishes since $\varepsilon_{ijj} = 0$, and thus

$$\hat{\mathbf{L}}\cdot\frac{\hat{\mathbf{r}}}{r} = 0. \tag{J.20}$$

Substitution of Eq. (J.17) and Eq. (J.20) in Eq. (J.2) finally yields the desired result:

$$\hat{\mathbf{L}}\cdot\hat{\mathbf{M}} = 0. \tag{J.21}$$

J.2 THE SCALAR PRODUCT $\hat{\mathbf{M}}^2$

The second identity of the LRL vector that we want to prove reads

$$\hat{\mathbf{M}}^2 = 2m\hat{\mathscr{H}}\left(\hat{\mathbf{L}}^2 + \hbar^2\right) + m^2k^2. \tag{J.22}$$

It is convenient, in this case, to use the alternative form of the LRL vector, as derived in Eq. (I.64). We can thus write

$$\begin{aligned}\hat{\mathbf{M}}^2 &= \left(\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right) - mk\frac{\hat{\mathbf{r}}}{r}\right)^2\\ &= \left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)^2 - mk\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)\cdot\frac{\hat{\mathbf{r}}}{r}\\ &\quad - mk\frac{\hat{\mathbf{r}}}{r}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right) + m^2k^2\frac{\hat{\mathbf{r}}^2}{r^2}.\end{aligned} \tag{J.23}$$

The first quadratic term in Eq. (J.23) can be written as

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)^2 = \left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)^2 - i\hbar\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\hat{\mathbf{p}} - i\hbar\hat{\mathbf{p}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right) - \hbar^2\hat{\mathbf{p}}^2. \tag{J.24}$$

Let us evaluate each of the four terms individually. The first term yields

$$\begin{aligned}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)^2 &= \sum_{i=1}^{3}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i\\ &= \sum_{ijklm}\varepsilon_{ijk}\varepsilon_{ilm}\hat{p}_j\hat{L}_k\hat{p}_l\hat{L}_m,\end{aligned} \tag{J.25}$$

where we have used the tensorial notation of Eq. (K.21) for the cross product of $\hat{\mathbf{p}}$ and $\hat{\mathbf{L}}$. The product of Levi-Civita symbols $\varepsilon_{ijk}\varepsilon_{ilm}$ can now be written in terms of Kronecker deltas with the help of the contracted epsilon identity from Eq. (K.5):

$$\begin{aligned}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)^2 &= \sum_{jklm}\left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right)\hat{p}_j\hat{L}_k\hat{p}_l\hat{L}_m\\ &= \sum_{jk}\left(\hat{p}_j\hat{L}_k\hat{p}_j\hat{L}_k - \hat{p}_j\hat{L}_k\hat{p}_k\hat{L}_j\right).\end{aligned} \tag{J.26}$$

From the commutation relations $\left[\hat{L}_k, \hat{p}_j\right] = i\hbar\varepsilon_{kjl}\hat{p}_l$, as derived in Eq. (I.20), we obtain for the first term

$$\sum_{jk}\hat{p}_j\left(\hat{L}_k\hat{p}_j\right)\hat{L}_k = \sum_{jk}\hat{p}_j^2\hat{L}_k^2 + i\hbar\sum_{jkl}\varepsilon_{kjl}\hat{p}_j\hat{p}_l\hat{L}_k = \hat{\mathbf{p}}^2\hat{\mathbf{L}}^2, \tag{J.27}$$

since $\sum_{j,l=1}^{3}\varepsilon_{kjl}\hat{p}_j\hat{p}_l = 0$ for all k. The second term in Eq. (J.26) can be expanded by writing $\hat{L}_k$ in tensor notation:

$$\begin{aligned}\sum_{jk}\hat{p}_j\hat{L}_k\hat{p}_k\hat{L}_j &= \sum_{jk}\hat{p}_j\left(\hat{\mathbf{r}}\times\hat{\mathbf{p}}\right)_k\hat{p}_k\hat{L}_j \\ &= \sum_{jklm}\varepsilon_{klm}\hat{p}_j\hat{r}_l\hat{p}_m\hat{p}_k\hat{L}_j = 0,\end{aligned} \tag{J.28}$$

which is seen to vanish as a consequence of the fact that $\sum_{k,m=1}^{3}\varepsilon_{klm}\hat{p}_m\hat{p}_k = 0$ for all l, thus yielding the overall result for Eq. (J.26):

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)^2 = \hat{\mathbf{p}}^2\hat{\mathbf{L}}^2. \tag{J.29}$$

Next, let us evaluate a part of the second term of Eq. (J.24):

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\hat{\mathbf{p}} = \sum_{i=1}^{3}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i\hat{p}_i = \sum_{ijk}\varepsilon_{ijk}\hat{p}_j\hat{L}_k\hat{p}_i. \tag{J.30}$$

Following $\left[\hat{L}_k,\hat{p}_i\right] = i\hbar\varepsilon_{kil}\hat{p}_l$, this becomes

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\hat{\mathbf{p}} = \sum_{ijk}\varepsilon_{ijk}\hat{p}_j\hat{p}_i\hat{L}_k + i\hbar\sum_{ijkl}\varepsilon_{ijk}\varepsilon_{kil}\hat{p}_j\hat{p}_l. \tag{J.31}$$

The first term vanishes since $\sum_{i,j=1}^{3}\varepsilon_{ijk}\hat{p}_j\hat{p}_i = 0$ for all k. The product of Levi-Civita symbols $\varepsilon_{ijk}\varepsilon_{kil}$ in the second term can be rewritten using the second contracted epsilon identity from Eq. (K.6),

$$\varepsilon_{ijk}\varepsilon_{kil} = \varepsilon_{kij}\varepsilon_{kil} = 2\delta_{jl}, \tag{J.32}$$

yielding

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\hat{\mathbf{p}} = 2i\hbar\sum_{jl}\delta_{jl}\hat{p}_j\hat{p}_l = 2i\hbar\sum_{j}\hat{p}_j\hat{p}_j = 2i\hbar\hat{\mathbf{p}}^2. \tag{J.33}$$

The third term in Eq. (J.24) can be evaluated in much the same vein. That is,

$$\hat{\mathbf{p}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right) = \sum_{i=1}^{3}\hat{p}_i\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i = \sum_{ijk}\varepsilon_{ijk}\hat{p}_i\hat{p}_j\hat{L}_k = 0, \tag{J.34}$$

which is seen to vanish in view of the fact that $\sum_{i,j=1}^{3}\varepsilon_{ijk}\hat{p}_i\hat{p}_j = 0$ for all k. After collecting all the partial results in Eq. (J.29), Eq. (J.33), and Eq. (J.34), we obtain for Eq. (J.24) the following result:

$$\begin{aligned}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}} - i\hbar\hat{\mathbf{p}}\right)^2 &= \left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)^2 - i\hbar\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\hat{\mathbf{p}} - i\hbar\hat{\mathbf{p}}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right) - \hbar^2\hat{\mathbf{p}}^2 \\ &= \hat{\mathbf{p}}^2\hat{\mathbf{L}}^2 + 2\hbar^2\hat{\mathbf{p}}^2 - \hbar^2\hat{\mathbf{p}}^2 \\ &= \hat{\mathbf{p}}^2\left(\hat{\mathbf{L}}^2 + \hbar^2\right).\end{aligned} \tag{J.35}$$

The second term to be evaluated in Eq. (J.23) yields

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\frac{\hat{\mathbf{r}}}{r} = \sum_{i=1}^{3}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i\frac{\hat{r}_i}{r} = \sum_{ijk}\varepsilon_{ijk}\hat{p}_j\hat{L}_k\frac{\hat{r}_i}{r}. \tag{J.36}$$

From the commutation relations $\left[\hat{L}_k,\hat{r}_i/r\right] = i\hbar\varepsilon_{kil}\hat{r}_l/r$, as derived in Eq. (I.57), we obtain

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\frac{\hat{\mathbf{r}}}{r} = \sum_{ijk}\varepsilon_{ijk}\hat{p}_j\frac{\hat{r}_i}{r}\hat{L}_k + i\hbar\sum_{ijkl}\varepsilon_{ijk}\varepsilon_{kil}\hat{p}_j\frac{\hat{r}_l}{r}. \tag{J.37}$$

The first term can be evaluated further with the help of the commutation relations $\left[\hat{p}_j, \hat{r}_i/r\right] = -i\hbar\left(\delta_{ij}1/r - \hat{r}_i\hat{r}_j/r^3\right)$ (see Eq. (I.41)):

$$\sum_{ijk}\varepsilon_{ijk}\hat{p}_j\frac{\hat{r}_i}{r}\hat{L}_k=\sum_{ijk}\varepsilon_{ijk}\frac{\hat{r}_i}{r}\hat{p}_j\hat{L}_k-i\hbar\sum_{ijk}\varepsilon_{ijk}\left(\delta_{ij}\frac{1}{r}-\frac{\hat{r}_i\hat{r}_j}{r^3}\right)\hat{L}_k. \tag{J.38}$$

The second term in Eq. (J.38) vanishes completely since $\sum_{ijk}\varepsilon_{ijk}\delta_{ij}=\sum_{ik}\varepsilon_{iik}=0$ and $\sum_{i,j=1}^{3}\varepsilon_{ijk}\hat{r}_i\hat{r}_j=0$ for all k. The first term can be rewritten as

$$\sum_{ijk}\varepsilon_{ijk}\frac{\hat{r}_i}{r}\hat{p}_j\hat{L}_k=\frac{1}{r}\sum_{k}\left(\hat{\mathbf{r}}\times\hat{\mathbf{p}}\right)_k\hat{L}_k=\frac{1}{r}\sum_{k}\hat{L}_k\hat{L}_k=\frac{1}{r}\hat{\mathbf{L}}^2. \tag{J.39}$$

Let us now compute the second term of Eq. (J.37). The product of Levi-Civita symbols $\varepsilon_{ijk}\varepsilon_{kil}$ can be reduced to $2\delta_{jl}$ by Eq. (J.32), yielding

$$\sum_{ijkl}\varepsilon_{ijk}\varepsilon_{kil}\hat{p}_j\frac{\hat{r}_l}{r}=\sum_{jl}2\delta_{jl}\hat{p}_j\frac{\hat{r}_l}{r}=\sum_{j}2\hat{p}_j\frac{\hat{r}_j}{r}=2\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}. \tag{J.40}$$

Substitution of Eq. (J.38), Eq. (J.39), and Eq. (J.40) into Eq. (J.37) finally yields

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)\cdot\frac{\hat{\mathbf{r}}}{r}=\frac{1}{r}\hat{\mathbf{L}}^2+2i\hbar\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}. \tag{J.41}$$

The third and last term to be evaluated in Eq. (J.23) contains the product

$$\frac{\hat{\mathbf{r}}}{r}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)=\sum_{i=1}^{3}\frac{\hat{r}_i}{r}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}\right)_i=\sum_{ijk}\varepsilon_{ijk}\frac{\hat{r}_i}{r}\hat{p}_j\hat{L}_k=\frac{1}{r}\hat{\mathbf{L}}^2, \tag{J.42}$$

by Eq. (J.39). Combining the results obtained in Eq. (J.41) and Eq. (J.42) enables us to rewrite the second and third terms of Eq. (J.23) as

$$\begin{aligned}\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)\cdot\frac{\hat{\mathbf{r}}}{r}+\frac{\hat{\mathbf{r}}}{r}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)&=\frac{1}{r}\hat{\mathbf{L}}^2+2i\hbar\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}-i\hbar\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}\\&\quad+\frac{1}{r}\hat{\mathbf{L}}^2-i\hbar\frac{\hat{\mathbf{r}}}{r}\cdot\hat{\mathbf{p}}\\&=\frac{2}{r}\hat{\mathbf{L}}^2+i\hbar\left(\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}-\frac{\hat{\mathbf{r}}}{r}\cdot\hat{\mathbf{p}}\right).\end{aligned} \tag{J.43}$$

The second term can be written as

$$\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}-\frac{\hat{\mathbf{r}}}{r}\cdot\hat{\mathbf{p}}=\sum_{i=1}^{3}\left(\hat{p}_i\frac{\hat{r}_i}{r}-\frac{\hat{r}_i}{r}\hat{p}_i\right)=\sum_{i=1}^{3}\left[\hat{p}_i,\frac{\hat{r}_i}{r}\right], \tag{J.44}$$

and with the help of $\left[\hat{p}_i,\frac{\hat{r}_i}{r}\right]=-i\hbar\left(\delta_{ii}\frac{1}{r}-\frac{\hat{r}_i\hat{r}_i}{r^3}\right)$, this becomes

$$\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}-\frac{\hat{\mathbf{r}}}{r}\cdot\hat{\mathbf{p}}=-i\hbar\sum_{i=1}^{3}\left(\frac{1}{r}-\frac{\hat{r}_i\hat{r}_i}{r^3}\right)=-i\hbar\left(\frac{3}{r}-\frac{\hat{\mathbf{r}}^2}{r^3}\right)=-2i\hbar\frac{1}{r}. \tag{J.45}$$

We then obtain for Eq. (J.43):

$$\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)\cdot\frac{\hat{\mathbf{r}}}{r}+\frac{\hat{\mathbf{r}}}{r}\cdot\left(\hat{\mathbf{p}}\times\hat{\mathbf{L}}-i\hbar\hat{\mathbf{p}}\right)=\frac{2}{r}\left(\hat{\mathbf{L}}^2+\hbar^2\right). \tag{J.46}$$

With the results obtained in Eq. (J.35) and Eq. (J.46), Eq. (J.23) finally yields the desired result:

$$\begin{aligned}\hat{\mathbf{M}}^2&=\hat{\mathbf{p}}^2\left(\hat{\mathbf{L}}^2+\hbar^2\right)-2m\frac{k}{r}\left(\hat{\mathbf{L}}^2+\hbar^2\right)+m^2k^2\\&=2m\left(\frac{\hat{\mathbf{p}}^2}{2m}-\frac{k}{r}\right)\left(\hat{\mathbf{L}}^2+\hbar^2\right)+m^2k^2\\\hat{\mathbf{M}}^2&=2m\mathscr{H}\left(\hat{\mathbf{L}}^2+\hbar^2\right)+m^2k^2.\end{aligned} \tag{J.47}$$

Appendix K
Indicial notation

The aim of this appendix is to provide a short introduction to the *indicial notation* in vector algebra, also commonly known as *tensor notation* or *subscript notation*. Central within this notational system are the two key symbols δ_{ij} and ε_{ijk}, which are defined first.

K.1 THE KRONECKER DELTA δ_{ij}

The *Kronecker delta*, denoted δ_{ij}, is defined as follows:

$$\delta_{ij} = \begin{cases} +1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}, \tag{K.1}$$

where i and j are dummy indices that can take on any value from 1 to 3 when working in three-dimensional space $\mathbb{R}^3$. Note also that

$$\delta_{ij} = \delta_{ji}. \tag{K.2}$$

The value of δ_{ii} can now be determined. Recall that, according to Einstein's summation convention, repeated indices imply summation over the range of possible values of these indices. In our case, we need to sum over all values of i from 1 to 3—that is,

$$\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 1 + 1 + 1 = 3. \tag{K.3}$$

K.2 THE LEVI-CIVITA TENSOR ε_{ijk}

Another useful indicial notation goes by the name *Levi-Civita symbol*, in honor of the Italian mathematician and physicist Tullio Levi-Civita (1873–1941). The elements of this *antisymmetric three-dimensional tensor* are defined by the permutation symbol ε_{ijk}, where

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) = (1,2,3),\ (2,3,1) \text{ or } (3,1,2) \\ -1 & \text{if } (i,j,k) = (1,3,2),\ (2,1,3) \text{ or } (3,2,1) \\ 0 & \text{if } i = j,\ j = k \text{ or } k = i \end{cases}. \tag{K.4}$$

K.2.1 Contracted epsilon identities

In this section, we will derive three useful identities that relate the Kronecker delta with the Levi-Civita symbol. These are known as the *contracted epsilon identities*:

$$\varepsilon_{ijk}\varepsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}, \tag{K.5}$$

$$\varepsilon_{ijk}\varepsilon_{ijn} = 2\delta_{kn}, \tag{K.6}$$

$$\varepsilon_{ijk}\varepsilon_{ijk} = 6, \tag{K.7}$$

where i, j, k, m, and n take any value from 1 to 3. As another reminder, Einstein summation over repeated indices is implied. In the first identity, for instance, the index i is repeated and is thought of as being summed over–that is,

$$\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}. \tag{K.8}$$

To prove these identities, we start by expressing the general product of two Levi-Civita symbols ε_{ijk} and ε_{lmn} as a function of Kronecker deltas:

$$\begin{aligned}\varepsilon_{ijk}\varepsilon_{lmn} &= \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \\ &\quad - \delta_{im}\delta_{jl}\delta_{kn} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{il}\delta_{jn}\delta_{km},\end{aligned} \tag{K.9}$$

the validity of which can be verified by explicit computation. Rewriting Eq. (K.9) as

$$\delta_{il}\left(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\right) + \delta_{im}\left(\delta_{jn}\delta_{kl} - \delta_{jl}\delta_{kn}\right) + \delta_{in}\left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right), \tag{K.10}$$

yields the determinental expression

$$\varepsilon_{ijk}\varepsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}. \tag{K.11}$$

Now, for the special case where $i = l$, Eq. (K.9) reads

$$\begin{aligned}\varepsilon_{ijk}\varepsilon_{imn} &= \delta_{ii}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{ki} + \delta_{in}\delta_{ji}\delta_{km} \\ &\quad - \delta_{im}\delta_{ji}\delta_{kn} - \delta_{in}\delta_{jm}\delta_{ki} - \delta_{ii}\delta_{jn}\delta_{km},\end{aligned} \tag{K.12}$$

which can be written as

$$\begin{aligned}\varepsilon_{ijk}\varepsilon_{imn} &= \delta_{ii}\left(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\right) + \delta_{im}\delta_{jn}\delta_{ki} + \delta_{in}\delta_{ji}\delta_{km} \\ &\quad - \delta_{im}\delta_{ji}\delta_{kn} - \delta_{in}\delta_{jm}\delta_{ki} \\ &= 3\left(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\right) + \delta_{km}\delta_{jn} + \delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn} - \delta_{kn}\delta_{jm} \\ &= \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km},\end{aligned} \tag{K.13}$$

thus proving the first identity in Eq. (K.5). Notice that the first delta in the right-hand side of Eq. (K.13) contains the *inner indices* from both epsilons, whereas the second delta uses both *outer indices.* The third delta, on the other hand, contains the inner index from the first epsilon and the outer index from the second epsilon—and conversely for the fourth delta. The order of the indices on the right-hand side of Eq. (K.13) can thus be remembered by:

$$\text{inner-inner}\quad\text{outer-outer}\quad-\quad\text{inner-outer}\quad\text{outer-inner}. \tag{K.14}$$

Of course, the order of the deltas within each term can be reversed, as can the order of the indices within each delta.

The final two identities are now easily obtained from the first one. Setting $j = m$ in Eq. (K.5) yields the second identity in Eq. (K.6):

$$\begin{aligned}\varepsilon_{ijk}\varepsilon_{ijn} &= \delta_{jj}\delta_{kn} - \delta_{jn}\delta_{kj} \\ &= 3\delta_{kn} - \delta_{kn} \\ &= 2\delta_{kn}.\end{aligned} \tag{K.15}$$

The third identity in Eq. (K.7) is finally obtained by setting $k = n$ in Eq. (K.6):

$$\varepsilon_{ijk}\varepsilon_{ijk} = 2\delta_{kk} = 2 \cdot 3 = 6. \tag{K.16}$$

K.2.2 Indicial notation for cross products

The cross product $\mathbf{a} \times \mathbf{b}$ of two vectors $\mathbf{a} = \{a_1, a_2, a_3\}$ and $\mathbf{b} = \{b_1, b_2, b_3\}$ can also be represented with the help of the Levi-Civita symbol. Given the determinental form of the cross product (see Appendix B)

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}, \tag{K.17}$$

the components of $\mathbf{a} \times \mathbf{b}$ can be written as

$$(\mathbf{a} \times \mathbf{b})_1 = a_2 b_3 - a_3 b_2; \tag{K.18}$$

$$(\mathbf{a} \times \mathbf{b})_2 = a_3 b_1 - a_1 b_3; \tag{K.19}$$

$$(\mathbf{a} \times \mathbf{b})_3 = a_1 b_2 - a_2 b_1. \tag{K.20}$$

Using the Levi-Civita symbol, this can be expressed more simply as

$$(\mathbf{a} \times \mathbf{b})_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} a_j b_k. \tag{K.21}$$

Finally, using Einstein's summation convention, this becomes

$$(\mathbf{a} \times \mathbf{b})_i = \varepsilon_{ijk} a_j b_k. \tag{K.22}$$

Appendix L
$\hat{S}$-transform of the $\mathfrak{so}(4,2)$ algebra

The derivations in this appendix are no longer based on the functional form of the operators, but on their actions in a basis set of $|nlm\rangle$ orbitals. For convenience, the actions of the $\mathfrak{so}(4,2)$ generators on a ket $|nlm\rangle$ are grouped here together:

$$
\begin{aligned}
\hat{L}_3\,|nlm\rangle &= m\,|nlm\rangle,\\
\hat{L}_\pm\,|nlm\rangle &= \omega^{l}_{\pm m}\,|nl\,(m\pm 1)\rangle,\\
\hat{A}_3\,|nlm\rangle &= \alpha^{l}_{m} c^{n}_{l}\,|n\,(l-1)\,m\rangle\\
&\quad +\alpha^{l+1}_{m}\,c^{n}_{l+1}\,|n\,(l+1)\,m\rangle,\\
\hat{A}_\pm\,|nlm\rangle &= \pm\beta^{l-1}_{\pm m} c^{n}_{l}\,|n\,(l-1)\,(m\pm 1)\rangle\\
&\quad \mp\gamma^{l+1}_{\pm m}\,c^{n}_{l+1}\,|n\,(l+1)\,(m\pm 1)\rangle,\\
\hat{B}_3\,|nlm\rangle &= \alpha^{l}_{m} u^{n}_{l}\,|(n-1)\,(l-1)\,m\rangle\\
&\quad +\alpha^{l}_{m} v^{n}_{l}\,|(n+1)\,(l-1)\,m\rangle\\
&\quad +\alpha^{l+1}_{m}\,v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,m\rangle\\
&\quad +\alpha^{l+1}_{m}\,u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,m\rangle,\\
\hat{B}_\pm\,|nlm\rangle &= \pm\beta^{l-1}_{\pm m} u^{n}_{l}\,|(n-1)\,(l-1)\,(m\pm 1)\rangle\\
&\quad \pm\beta^{l-1}_{\pm m} v^{n}_{l}\,|(n+1)\,(l-1)\,(m\pm 1)\rangle\\
&\quad \mp\gamma^{l+1}_{\pm m}\,v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,(m\pm 1)\rangle\\
&\quad \mp\gamma^{l+1}_{\pm m}\,u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,(m\pm 1)\rangle,\\
\hat{\Gamma}_3\,|nlm\rangle &= -i\alpha^{l}_{m} u^{n}_{l}\,|(n-1)\,(l-1)\,m\rangle\\
&\quad +i\alpha^{l}_{m} v^{n}_{l}\,|(n+1)\,(l-1)\,m\rangle\\
&\quad -i\alpha^{l+1}_{m}\,v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,m\rangle\\
&\quad +i\alpha^{l+1}_{m}\,u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,m\rangle,\\
\hat{\Gamma}_\pm\,|nlm\rangle &= \mp i\beta^{l-1}_{\pm m} u^{n}_{l}\,|(n-1)\,(l-1)\,(m\pm 1)\rangle\\
&\quad \pm i\beta^{l-1}_{\pm m} v^{n}_{l}\,|(n+1)\,(l-1)\,(m\pm 1)\rangle\\
&\quad \pm i\gamma^{l+1}_{\pm m}\,v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,(m\pm 1)\rangle\\
&\quad \mp i\gamma^{l+1}_{\pm m}\,u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,(m\pm 1)\rangle.\\
\hat{Q}_3\,|nlm\rangle &= n\,|nlm\rangle,\\
\hat{Q}_\pm\,|nlm\rangle &= \omega^{\pm n}_{l}\,|(n\pm 1)\,lm\rangle,
\end{aligned}
\tag{L.1}
$$

$$\text{where } \alpha^l_m = \sqrt{(l-m)\,(l+m)}; \tag{L.2}$$

$$\beta^l_m = \sqrt{(l-m+1)\,(l-m)}; \tag{L.3}$$

$$\gamma^l_m = \sqrt{(l+m+1)\,(l+m)}; \tag{L.4}$$

$$\omega^l_m = \sqrt{(l-m)\,(l+m+1)}; \tag{L.5}$$

$$c^n_l = \sqrt{\frac{(n-l)\,(n+l)}{(2l-1)\,(2l+1)}}; \tag{L.6}$$

$$u^n_l = \frac{1}{2}\sqrt{\frac{(n+l-1)\,(n+l)}{(2l-1)\,(2l+1)}}; \tag{L.7}$$

$$v^n_l = \frac{1}{2}\sqrt{\frac{(n-l)\,(n-l+1)}{(2l-1)\,(2l+1)}}. \tag{L.8}$$

The $\hat{S}$ operator plays the role of a *transformation operator*, under which the generators of $\mathfrak{so}(4,2)$ transform into new operators. Since $\hat{S}$ commutes with $\hat{L}_i$ and $\hat{Q}_i$ $(i = 1 \rightarrow 3)$, the operators $\hat{L}_i$ and $\hat{Q}_i$ remain the same. The other operators of $\mathfrak{so}(4,2)$ are transformed into primed equivalents, which are defined as follows:

$$\left[\hat{S}, \hat{A}_i\right] = \hat{A}'_i, \tag{L.9}$$

$$\left[\hat{S}, \hat{B}_i\right] = \hat{B}'_i, \tag{L.10}$$

$$\left[\hat{S}, \hat{\Gamma}_i\right] = \hat{\Gamma}'_i, \tag{L.11}$$

with $i = 3, +, -$. Their actions on a ket $|nlm\rangle$ are then given by

$$\begin{aligned}
\hat{A}'_3\,|nlm\rangle &= -\alpha^l_m c^n_l\,|n\,(l-1)\,m\rangle \\
&\quad + \alpha^{l+1}_m c^n_{l+1}\,|n\,(l+1)\,m\rangle; \\
\hat{A}'_\pm\,|nlm\rangle &= \mp\beta^{l-1}_{\pm m} c^n_l\,|n\,(l-1)\,(m\pm 1)\rangle \\
&\quad \mp \gamma^{l+1}_{\pm m} c^n_{l+1}\,|n\,(l+1)\,(m\pm 1)\rangle; \\
\hat{B}'_3\,|nlm\rangle &= -\alpha^l_m u^n_l\,|(n-1)\,(l-1)\,m\rangle \\
&\quad - \alpha^l_m v^n_l\,|(n+1)\,(l-1)\,m\rangle \\
&\quad + \alpha^{l+1}_m v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,m\rangle \\
&\quad + \alpha^{l+1}_m u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,m\rangle; \\
\hat{B}'_\pm\,|nlm\rangle &= \mp\beta^{l-1}_{\pm m} u^n_l\,|(n-1)\,(l-1)\,(m\pm 1)\rangle \\
&\quad \mp \beta^{l-1}_{\pm m} v^n_l\,|(n+1)\,(l-1)\,(m\pm 1)\rangle \\
&\quad \mp \gamma^{l+1}_{\pm m} v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,(m\pm 1)\rangle \\
&\quad \mp \gamma^{l+1}_{\pm m} u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,(m\pm 1)\rangle; \\
\hat{\Gamma}'_3\,|nlm\rangle &= +i\alpha^l_m u^n_l\,|(n-1)\,(l-1)\,m\rangle \\
&\quad - i\alpha^l_m v^n_l\,|(n+1)\,(l-1)\,m\rangle \\
&\quad - i\alpha^{l+1}_m v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,m\rangle \\
&\quad + i\alpha^{l+1}_m u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,m\rangle;
\end{aligned} \tag{L.12}$$

$$\begin{aligned}\hat{\Gamma}'_{\pm}\,|nlm\rangle = &\pm i\beta^{l-1}_{\pm m}u^{n}_{l}\,|(n-1)\,(l-1)\,(m\pm 1)\rangle\\ &\mp i\beta^{l-1}_{\pm m}v^{n}_{l}\,|(n+1)\,(l-1)\,(m\pm 1)\rangle\\ &\pm i\gamma^{l+1}_{\pm m}v^{n-1}_{l+1}\,|(n-1)\,(l+1)\,(m\pm 1)\rangle\\ &\mp i\gamma^{l+1}_{\pm m}u^{n+1}_{l+1}\,|(n+1)\,(l+1)\,(m\pm 1)\rangle.\end{aligned}$$

These actions are identical to the original ones effected under $\hat{A}_i$, $\hat{B}_i$, and $\hat{\Gamma}_i$, except for sign changes whenever l is lowered to $l-1$. This has an important consequence, though. Consider a matrix element in one of these operators—say, $\hat{B}'_3$—between states with different l-values. In the matrix elements where the l-value of the bra is lower than the l-value of the ket, the primed operator lowers the angular momentum of the ket function on which it is acting. Hence, it introduces an extra minus sign compared with the action of its unprimed analogue, $\hat{B}_3$. However, for the conjugate element, the change of angular momentum is in the opposite sense, from $l-1$ back to l. Hence, the angular momentum is raised and there is no extra minus sign. From Eq. (L.12), this is, indeed, verified, for example, in the case of the $\hat{B}'_3$ operator:

$$\begin{aligned}\langle (n-1)(l-1)(m)\,|\hat{B}'_3|\,nlm\rangle &= -\alpha^{l}_{m}u^{n}_{l}\\ \langle \hat{B}'_3(n-1)(l-1)(m)\,|\,nlm\rangle &= \alpha^{l}_{m}u^{n}_{l}.\end{aligned}\tag{L.13}$$

Since the matrix elements are real and have opposite sign, the result implies that the $\hat{B}'$ operators are *anti-Hermitian*, and similarly for the $\hat{A}'$ operators. For the $\hat{\Gamma}'$ operators, the equivalent result reads

$$\begin{aligned}\langle (n-1)(l-1)(m)\,|\hat{\Gamma}'_3|\,nlm\rangle &= i\alpha^{l}_{m}u^{n}_{l}\\ \langle nlm|\hat{\Gamma}'_3|\,(n-1)(l-1)(m)\rangle &= i\alpha^{l}_{m}u^{n}_{l}.\end{aligned}\tag{L.14}$$

In this case, the matrix elements are equal, with the same sign. However, they are both entirely imaginary, thus indicating that this operator is anti-Hermitian too.

To ensure that the transformed operators obey the proper $\mathfrak{so}(4,2)$ commutation relations, we must multiply them by the imaginary unit, which renders Hermiticity. Let us denote these Hermitian transforms as follows:

$$\begin{aligned}\hat{\mathcal{A}}_k &= i\hat{A}'_k\\ \hat{\mathcal{B}}_k &= i\hat{B}'_k\\ \hat{\mathcal{C}}_k &= i\hat{\Gamma}'_k.\end{aligned}\tag{L.15}$$

The newly obtained operators $\hat{\mathcal{A}}_k$, $\hat{\mathcal{B}}_k$, and $\hat{\mathcal{C}}_k$, along with the original $\hat{L}_i$ and $\hat{Q}_i$, form a reflected $\mathfrak{so}'(4,2)$ Lie algebra. It can be verified that the commutation relations in the transformed algebra mimic precisely the ones in the original algebra. Let us examine this by working out a few commutators. For the commutator $\left[\hat{\mathcal{A}}_{\pm},\hat{\mathcal{A}}_3\right]$, the consecutive actions of $\hat{\mathcal{A}}_3$ and $\hat{\mathcal{A}}_{\pm}$ are

$$\begin{aligned}\hat{\mathcal{A}}_{\pm}\hat{\mathcal{A}}_3\,|nlm\rangle = &\,\lambda_{\pm 1}\,|(n)(l-2)(m\pm 1)\rangle\\ &+\lambda_{\pm 2}\,|nl(m\pm 1)\rangle\\ &+\lambda_{\pm 3}\,|(n)(l+2)(m\pm 1)\rangle;\end{aligned}\tag{L.16}$$

$$\begin{aligned}\hat{\mathcal{A}}_3\hat{\mathcal{A}}_{\pm}\,|nlm\rangle = &\,\lambda_{\pm 4}\,|(n)(l-2)(m\pm 1)\rangle\\ &+\lambda_{\pm 5}\,|nl(m\pm 1)\rangle\\ &+\lambda_{\pm 6}\,|(n)(l+2)(m\pm 1)\rangle.\end{aligned}\tag{L.17}$$

The coefficients in this equation can be identified with the help of Eq. (L.12):

$$\begin{aligned}
\lambda_{\pm 1} - \lambda_{\pm 4} &= \mp c_l^n c_{l-1}^n \left(\alpha_m^l \beta_{\pm m}^{l-2} - \alpha_{m\pm 1}^{l-1} \beta_{\pm m}^{l-1} \right) = 0 \\
\lambda_{\pm 3} - \lambda_{\pm 6} &= \pm c_{l+1}^n c_{l+2}^n \left(\alpha_m^{l+1} \gamma_{\pm m}^{l+2} - \alpha_{m\pm 1}^{l+2} \gamma_{\pm m}^{l+1} \right) = 0 \\
\lambda_{\pm 2} - \lambda_{\pm 5} &= \mp \left(c_l^n\right)^2 \left(\alpha_m^l \gamma_{\pm m}^l + \alpha_{m\pm 1}^l \beta_{\pm m}^{l-1} \right) \pm \left(c_{l+1}^n\right)^2 \left(\alpha_m^{l+1} \beta_{\pm m}^l + \alpha_{m\pm 1}^{l+1} \gamma_{\pm m}^{l+1} \right) \\
&= \mp \omega_{\pm m}^l \left(\frac{n^2 - l^2}{2l+1} - \frac{n^2 - (l+1)^2}{2l+1} \right) \\
&= \mp \omega_{\pm m}^l .
\end{aligned} \tag{L.18}$$

Hence, the commutator is expressed as

$$\left[\hat{\mathcal{A}}_\pm , \hat{\mathcal{A}}_3 \right] = \mp \hat{L}_\pm . \tag{L.19}$$

This result is identical to the commutation relations between the components of the LRL vector:

$$\left[\hat{A}_\pm , \hat{A}_3 \right] = \mp \hat{L}_\pm . \tag{L.20}$$

As a final example, let us compute the commutator between $\hat{\mathcal{C}}_+$ and $\hat{\mathcal{C}}_3$. These operators can only change the sum of $n + l$ by an even number, and their product gives rise to nine different combinations:

$$\begin{pmatrix} n \\ l \end{pmatrix} \rightarrow \begin{pmatrix} n\pm 2 \\ l\pm 2 \end{pmatrix}, \begin{pmatrix} n\pm 2 \\ l\mp 2 \end{pmatrix}, \begin{pmatrix} n\pm 2 \\ l \end{pmatrix}, \begin{pmatrix} n \\ l\pm 2 \end{pmatrix}, \begin{pmatrix} n \\ l \end{pmatrix}. \tag{L.21}$$

A lengthy calculation shows that all these contributions vanish, except for the diagonal element. The separate contributions involve summations over several paths. The simultaneous change of n and l by 2 units can be done only in one way, the raising or lowering of either n or l by 2 units can be achieved by two paths, and, finally, the diagonal element is the summation over four coupling paths. The four elements with single coupling paths are as follows:

$$\begin{aligned}
\left\langle \begin{matrix} n+2 \\ l+2 \\ m+1 \end{matrix} \middle| \left[\hat{\mathcal{C}}_+ , \hat{\mathcal{C}}_3 \right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= u_{l+1}^{n+1} u_{l+2}^{n+2} \left(\alpha_m^{l+1} \gamma_m^{l+2} - \alpha_{m+1}^{l+2} \gamma_m^{l+1} \right) = 0; \\
\left\langle \begin{matrix} n+2 \\ l-2 \\ m+1 \end{matrix} \middle| \left[\hat{\mathcal{C}}_+ , \hat{\mathcal{C}}_3 \right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= v_l^n v_{l-1}^{n+1} \left(\alpha_m^l \beta_m^{l-2} - \alpha_{m+1}^{l-1} \beta_m^{l-1} \right) = 0; \\
\left\langle \begin{matrix} n-2 \\ l+2 \\ m+1 \end{matrix} \middle| \left[\hat{\mathcal{C}}_+ , \hat{\mathcal{C}}_3 \right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= -v_{l+1}^{n-1} v_{l+2}^{n-2} \left(\alpha_m^{l+1} \gamma_m^{l+2} - \alpha_{m+1}^{l+2} \gamma_m^{l+1} \right) = 0; \\
\left\langle \begin{matrix} n-2 \\ l-2 \\ m+1 \end{matrix} \middle| \left[\hat{\mathcal{C}}_+ , \hat{\mathcal{C}}_3 \right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= u_l^n u_{l-1}^{n-1} \left(\alpha_m^l \beta_m^{l-2} - \alpha_{m+1}^{l-1} \beta_m^{l-1} \right) = 0.
\end{aligned} \tag{L.22}$$

For the matrix elements that involve two alternative pathways, we have

$$\begin{aligned}
\left\langle \begin{matrix} n+2 \\ l \\ m+1 \end{matrix} \middle| \left[\hat{C}_+,\hat{C}_3\right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= u_l^{n+2} v_l^n \left(\alpha_m^l \gamma_m^l + \alpha_{m+1}^l \beta_m^{l-1}\right) \\
&\quad - u_{l+1}^{n+1} v_{l+1}^{n+1} \left(\alpha_m^{l+1} \beta_m^l + \alpha_{m+1}^{l+1} \gamma_m^{l+1}\right) = 0; \\
\left\langle \begin{matrix} n-2 \\ l \\ m+1 \end{matrix} \middle| \left[\hat{C}_+,\hat{C}_3\right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= -u_{l+1}^{n-1} v_{l+1}^{n-1} \left(\alpha_m^{l+1} \beta_m^l + \alpha_{m+1}^{l+1} \gamma_m^{l+1}\right) \\
&\quad + u_l^n v_l^{n-2} \left(\alpha_m^l \gamma_m^l + \alpha_{m+1}^l \beta_m^{l-1}\right) = 0; \\
\left\langle \begin{matrix} n \\ l+2 \\ m+1 \end{matrix} \middle| \left[\hat{C}_+,\hat{C}_3\right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= u_{l+1}^{n+1} v_{l+2}^n \left(\alpha_m^{l+1} \gamma_m^{l+2} - \alpha_{m+1}^{l+2} \gamma_m^{l+1}\right) \\
&\quad + u_{l+2}^n v_{l+1}^{n-1} \left(\alpha_m^{l+1} \gamma_m^{l+2} - \alpha_{m+1}^{l+2} \gamma_m^{l+1}\right) = 0; \\
\left\langle \begin{matrix} n \\ l-2 \\ m+1 \end{matrix} \middle| \left[\hat{C}_+,\hat{C}_3\right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle &= -u_{l-1}^{n+1} v_l^n \left(\alpha_m^l \beta_m^{l-2} - \alpha_{m+1}^{l-1} \beta_m^{l-1}\right) \\
&\quad - u_l^n v_{l-1}^{n-1} \left(\alpha_m^l \beta_m^{l-2} - \alpha_{m+1}^{l-1} \beta_m^{l-1}\right) = 0.
\end{aligned} \tag{L.23}$$

The first of these zeros is a result of the following interesting sum rule:

$$u_l^{n+2} v_l^n \, (2l-1) = u_{l+1}^{n+1} v_{l+1}^{n+1} \, (2l+3). \tag{L.24}$$

Finally, the diagonal element involves four coupling pathways, covering all combinations of changing both n and l by ±1. We have

$$\begin{aligned}
&\left\langle \begin{matrix} n \\ l \\ M+1 \end{matrix} \middle| \left[\hat{C}_+,\hat{C}_3\right] \middle| \begin{matrix} n \\ l \\ m \end{matrix} \right\rangle \\
&= \left(u_{l+1}^{n+1}\right)^2 \left(\alpha_m^{l+1} \beta_m^l + \alpha_{m+1}^{l+1} \gamma_m^{l+1}\right) - \left(v_l^n\right)^2 \left(\alpha_m^l \gamma_m^l + \alpha_{m+1}^l \beta_m^{l-1}\right) \\
&\quad + \left(v_{l+1}^{n-1}\right)^2 \left(\alpha_m^{l+1} \beta_m^l + \alpha_{m+1}^{l+1} \gamma_m^{l+1}\right) - \left(u_l^n\right)^2 \left(\alpha_m^l \gamma_m^l + \alpha_{m+1}^l \beta_m^{l-1}\right) \\
&= \omega_m^l.
\end{aligned} \tag{L.25}$$

Hence, the commutator reduces to

$$\left[\hat{C}_+,\hat{C}_3\right] = \hat{L}_+. \tag{L.26}$$

This result is a perfect match of the commutator between $\hat{\Gamma}_+$ and $\hat{\Gamma}_3$.

Index