College Physics
Textbook Equity Edition

Volume 3 of 3: Chapters 25 - 34

This multi-part Textbook Equity edition retains the original academic content as published by Openstax College, and under the terms of their Creative Commons license (CC-BY).

TextbookEquity.org
Provider of truly open textbooks

"College Physics" Textbook Equity Edition,
Vol 3 (of 3) License: CC-BY
Free pdf or low cost print at:
http://textbookequity.org/collegephysics_vol3
This is volume 3 of 3, covering chapters 25 - 34, plus the complete table of contents, appendices and index. Page numbers refer to the original 1269 page textbook published by OpenStax College as noted below. As a courtesy the original front matter, with some additional informate, and the preface are included in this edition.

Textbook Equity* (http://textbookequity.org) provides high quality and technically convenient academic content and free access to any learner at any location, at the minimum possible cost. Education of everyone is the only true means to eliminate poverty and oppression. Creating barriers to education, whether intentional or not, by policy or price, leaves millions of geniuses undiscovered. We work to get them discovered. (*fearlessly copy, print, remix)

Visit http://textbookequity.org for the full repository of truly open licensed textbooks (free pdf and inexpensive print)

Original Front Matter

OpenStax College Rice University
6100 Main Street MS-380 Houston, Texas 77005

To learn more about OpenStax College, visit http://openstaxcollege.org. Individual print copies and bulk orders can be purchased through our website.

© 2013 Rice University. Textbook content produced by OpenStax College is licensed under a Creative Commons Attribution 3.0 Unported License. Under this license, any user of this textbook or the textbook contents herein must provide proper attribution as follows:

- If you redistribute this textbook in a digital format (including but not limited to EPUB, PDF, and HTML), then you must retain on every page the following attribution:
  “Download for free at http://cnx.org/content/col11406/latest/.”
- If you redistribute this textbook in a print format, then you must include on every physical page the following attribution:
  “Download for free at http://cnx.org/content/col11406/latest/.”
- If you redistribute part of this textbook, then you must retain in every digital format page view (including but not limited to EPUB, PDF, and HTML) and on every physical printed page the following attribution:
  “Download for free at http://cnx.org/content/col11406/latest/.”
- If you use this textbook as a bibliographic reference, then you should cite it as follows: OpenStax College, College Physics. OpenStax College. 21 June 2012. <http://cnx.org/content/col11406/latest/>.

For questions regarding this licensing, please contact partners@openstaxcollege.org.

Trademarks

The OpenStax College name, OpenStax College logo, OpenStax College book covers, Connexions name, and Connexions logo are registered trademarks of Rice University. All rights reserved. Any of the trademarks, service marks, collective marks, design rights, or similar rights that are mentioned, used, or cited in OpenStax College, Connexions, or Connexions’ sites are the property of their respective owners.

ISBN-10 1938168003
ISBN-13 978-1-938168-00-0
Revision CP-1-004-DW

License: CC-BY
http://textbookequity.org/collegephysics_vol3
Textbook Equity* (textbookequity.org) provides high quality academic content and free access to any learner at any location, at the minimum possible cost. Education of everyone is the only true means to eliminate poverty and oppression. Creating barriers to education, whether intentional or not, by policy or price, leaves millions of geniuses undiscovered. We want them discovered. “Fearlessly copy, print, remix.”\textsuperscript{tm}

OpenStax College

OpenStax College is a non-profit organization committed to improving student access to quality learning materials. Our free textbooks are developed and peer-reviewed by educators to ensure they are readable, accurate, and meet the scope and sequence requirements of modern college courses. Through our partnerships with companies and foundations committed to reducing costs for students, OpenStax College is working to improve access to higher education for all.

Connexions

The technology platform supporting OpenStax College is Connexions (http://cnx.org), one of the world’s first and largest open-education projects. Connexions provides students with free online and low-cost print editions of the OpenStax College library and provides instructors with tools to customize the content so that they can have the perfect book for their course.

Rice University

OpenStax College and Connexions are initiatives of Rice University. As a leading research university with a distinctive commitment to undergraduate education, Rice University aspires to path-breaking research, unsurpassed teaching, and contributions to the betterment of our world. It seeks to fulfill this mission by cultivating a diverse community of learning and discovery that produces leaders across the spectrum of human endeavor.

Foundation Support

OpenStax College is grateful for the tremendous support of our sponsors. Without their strong engagement, the goal of free access to high-quality textbooks would remain just a dream.

The William and Flora Hewlett Foundation has been making grants since 1967 to help solve social and environmental problems at home and around the world. The Foundation concentrates its resources on activities in education, the environment, global development and population, performing arts, and philanthropy, and makes grants to support disadvantaged communities in the San Francisco Bay Area.

Guided by the belief that every life has equal value, the Bill & Melinda Gates Foundation works to help all people lead healthy, productive lives. In developing countries, it focuses on improving people’s health with vaccines and other life-saving tools and giving them the chance to lift themselves out of hunger and extreme poverty. In the United States, it seeks to significantly improve education so that all young people have the opportunity to reach their full potential. Based in Seattle, Washington, the foundation is led by CEO Jeff Raikes and Co-chair William H. Gates Sr., under the direction of Bill and Melinda Gates and Warren Buffett.

Our mission at the Twenty Million Minds Foundation is to grow access and success by eliminating unnecessary hurdles to affordability. We support the creation, sharing, and proliferation of more effective, more affordable educational content by leveraging disruptive technologies, open educational resources, and new models for collaboration between for-profit, nonprofit, and public entities.

The Maxfield Foundation supports projects with potential for high impact in science, education, sustainability, and other areas of social importance.
Volume 1 Chapter Headings
1 Introduction: The Nature of Science and Physics
2 Kinematics
3 Two-Dimensional Kinematics
4 Dynamics: Force and Newton’s Laws of Motion
5 Further Applications of Newton’s Laws: Friction, Drag, and Elasticity
6 Uniform Circular Motion and Gravitation
7 Work, Energy, and Energy Resources
8 Linear Momentum and Collisions
9 Statics and Torque
10 Rotational Motion and Angular Momentum
11 Fluid Statics
12 Fluid Dynamics and Its Biological and Medical Applications
Appendices
A Atomic Masses
B Selected Radioactive Isotopes
C Useful Information
D Glossary of Key Symbols and Notation
Index (Vol 1 – 3)

Volume 2 Chapter Headings
13 Temperature, Kinetic Theory, and the Gas Laws
14 Heat and Heat Transfer Methods
15 Thermodynamics
16 Oscillatory Motion and Waves
17 Physics of Hearing
18 Electric Charge and Electric Field
19 Electric Potential and Electric Field
20 Electric Current, Resistance, and Ohm’s Law
21 Circuits, Bioelectricity, and DC Instruments
22 Magnetism
23 Electromagnetic Induction, AC Circuits, and Electrical Technologies
24 Electromagnetic Waves
Appendices
A Atomic Masses
B Selected Radioactive Isotopes
C Useful Information
D Glossary of Key Symbols and Notation
Index (Vol 1 – 3)

Volume 3 Chapter Headings
25 Geometric Optics
26 Vision and Optical Instruments
27 Wave Optics
28 Special Relativity
29 Introduction to Quantum Physics
30 Atomic Physics
31 Radioactivity and Nuclear Physics
32 Medical Applications of Nuclear Physics
33 Particle Physics
34 Frontiers of Physics
Appendices
A Atomic Masses
B Selected Radioactive Isotopes
C Useful Information
D Glossary of Key Symbols and Notation
Index (Vol 1 – 3)
# Table of Contents Vol 1 - 3

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>Introduction: The Nature of Science and Physics</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Physics: An Introduction</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Physical Quantities and Units</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Accuracy, Precision, and Significant Figures</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Approximation</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>Kinematics</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Displacement</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Vectors, Scalars, and Coordinate Systems</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Time, Velocity, and Speed</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Acceleration</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Motion Equations for Constant Acceleration in One Dimension</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Problem-Solving Basics for One-Dimensional Kinematics</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Falling Objects</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Graphical Analysis of One-Dimensional Motion</td>
<td>68</td>
</tr>
<tr>
<td>3</td>
<td>Two-Dimensional Kinematics</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Kinematics in Two Dimensions: An Introduction</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Vector Addition and Subtraction: Graphical Methods</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Vector Addition and Subtraction: Analytical Methods</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Projectile Motion</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>Addition of Velocities</td>
<td>108</td>
</tr>
<tr>
<td>4</td>
<td>Dynamics: Force and Newton’s Laws of Motion</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>Development of Force Concept</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>Newton’s First Law of Motion: Inertia</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>Newton’s Second Law of Motion: Concept of a System</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>Newton’s Third Law of Motion: Symmetry in Forces</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>Normal, Tension, and Other Examples of Forces</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Problem-Solving Strategies</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>Further Applications of Newton’s Laws of Motion</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>Extended Topic: The Four Basic Forces—An Introduction</td>
<td>152</td>
</tr>
<tr>
<td>5</td>
<td>Further Applications of Newton’s Laws: Friction, Drag, and Elasticity</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td>Friction</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td>Drag Forces</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>Elasticity: Stress and Strain</td>
<td>174</td>
</tr>
<tr>
<td>6</td>
<td>Uniform Circular Motion and Gravitation</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Rotation Angle and Angular Velocity</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Centripetal Acceleration</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>Centripetal Force</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>Fictitious Forces and Non-inertial Frames: The Coriolis Force</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>Newton’s Universal Law of Gravitation</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>Satellites and Kepler’s Laws: An Argument for Simplicity</td>
<td>209</td>
</tr>
<tr>
<td>7</td>
<td>Work, Energy, and Energy Resources</td>
<td>223</td>
</tr>
<tr>
<td></td>
<td>Work: The Scientific Definition</td>
<td>223</td>
</tr>
<tr>
<td></td>
<td>Kinetic Energy and the Work-Energy Theorem</td>
<td>226</td>
</tr>
<tr>
<td></td>
<td>Gravitational Potential Energy</td>
<td>230</td>
</tr>
<tr>
<td></td>
<td>Conservative Forces and Potential Energy</td>
<td>235</td>
</tr>
<tr>
<td></td>
<td>Nonconservative Forces</td>
<td>238</td>
</tr>
<tr>
<td></td>
<td>Conservation of Energy</td>
<td>242</td>
</tr>
<tr>
<td></td>
<td>Power</td>
<td>245</td>
</tr>
<tr>
<td></td>
<td>Work, Energy, and Power in Humans</td>
<td>249</td>
</tr>
<tr>
<td></td>
<td>World Energy Use</td>
<td>251</td>
</tr>
<tr>
<td>8</td>
<td>Linear Momentum and Collisions</td>
<td>263</td>
</tr>
<tr>
<td></td>
<td>Linear Momentum and Force</td>
<td>263</td>
</tr>
<tr>
<td></td>
<td>Impulse</td>
<td>265</td>
</tr>
<tr>
<td></td>
<td>Conservation of Momentum</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>Elastic Collisions in One Dimension</td>
<td>270</td>
</tr>
<tr>
<td></td>
<td>Inelastic Collisions in One Dimension</td>
<td>272</td>
</tr>
<tr>
<td></td>
<td>Collisions of Point Masses in Two Dimensions</td>
<td>276</td>
</tr>
<tr>
<td></td>
<td>Introduction to Rocket Propulsion</td>
<td>279</td>
</tr>
<tr>
<td>9</td>
<td>Statics and Torque</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>The First Condition for Equilibrium</td>
<td>292</td>
</tr>
<tr>
<td></td>
<td>The Second Condition for Equilibrium</td>
<td>293</td>
</tr>
<tr>
<td></td>
<td>Stability</td>
<td>297</td>
</tr>
<tr>
<td></td>
<td>Applications of Statics, Including Problem-Solving Strategies</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>Simple Machines</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>Forces and Torques in Muscles and Joints</td>
<td>306</td>
</tr>
<tr>
<td>10</td>
<td>Rotational Motion and Angular Momentum</td>
<td>319</td>
</tr>
<tr>
<td></td>
<td>Angular Acceleration</td>
<td>320</td>
</tr>
<tr>
<td></td>
<td>Kinematics of Rotational Motion</td>
<td>324</td>
</tr>
<tr>
<td></td>
<td>Dynamics of Rotational Motion: Rotational Inertia</td>
<td>328</td>
</tr>
<tr>
<td></td>
<td>Rotational Kinetic Energy: Work and Energy Revisited</td>
<td>331</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>11</td>
<td>Fluid Statics</td>
<td>359</td>
</tr>
<tr>
<td></td>
<td>What Is a Fluid?</td>
<td>359</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Pressure</td>
<td>362</td>
</tr>
<tr>
<td></td>
<td>Variation of Pressure with Depth in a Fluid</td>
<td>364</td>
</tr>
<tr>
<td></td>
<td>Pascal's Principle</td>
<td>367</td>
</tr>
<tr>
<td></td>
<td>Gauge Pressure, Absolute Pressure, and Pressure Measurement</td>
<td>370</td>
</tr>
<tr>
<td></td>
<td>Archimedes' Principle</td>
<td>373</td>
</tr>
<tr>
<td></td>
<td>Cohesion and Adhesion in Liquids: Surface Tension and Capillary Action</td>
<td>379</td>
</tr>
<tr>
<td></td>
<td>Pressures in the Body</td>
<td>386</td>
</tr>
<tr>
<td>12</td>
<td>Fluid Dynamics and Its Biological and Medical Applications</td>
<td>399</td>
</tr>
<tr>
<td></td>
<td>Flow Rate and Its Relation to Velocity</td>
<td>399</td>
</tr>
<tr>
<td></td>
<td>Bernoulli's Equation</td>
<td>402</td>
</tr>
<tr>
<td></td>
<td>The Most General Applications of Bernoulli's Equation</td>
<td>406</td>
</tr>
<tr>
<td></td>
<td>Viscosity and Laminar Flow, Poiseuille's Law</td>
<td>409</td>
</tr>
<tr>
<td></td>
<td>The Onset of Turbulence</td>
<td>415</td>
</tr>
<tr>
<td></td>
<td>Motion of an Object in a Viscous Fluid</td>
<td>416</td>
</tr>
<tr>
<td></td>
<td>Molecular Transport Phenomena: Diffusion, Osmosis, and Related Processes</td>
<td>418</td>
</tr>
<tr>
<td>13</td>
<td>Temperature, Kinetic Theory, and the Gas Laws</td>
<td>431</td>
</tr>
<tr>
<td></td>
<td>Temperature</td>
<td>431</td>
</tr>
<tr>
<td></td>
<td>Thermal Expansion of Solids and Liquids</td>
<td>438</td>
</tr>
<tr>
<td></td>
<td>The Ideal Gas Law</td>
<td>444</td>
</tr>
<tr>
<td></td>
<td>Kinetic Theory: Atomic and Molecular Explanation of Pressure and Temperature</td>
<td>449</td>
</tr>
<tr>
<td></td>
<td>Phase Changes</td>
<td>455</td>
</tr>
<tr>
<td></td>
<td>Humidity, Evaporation, and Boiling</td>
<td>460</td>
</tr>
<tr>
<td>14</td>
<td>Heat and Heat Transfer Methods</td>
<td>471</td>
</tr>
<tr>
<td></td>
<td>Heat</td>
<td>471</td>
</tr>
<tr>
<td></td>
<td>Temperature Change and Heat Capacity</td>
<td>473</td>
</tr>
<tr>
<td></td>
<td>Phase Change and Latent Heat</td>
<td>477</td>
</tr>
<tr>
<td></td>
<td>Heat Transfer Methods</td>
<td>482</td>
</tr>
<tr>
<td></td>
<td>Conduction</td>
<td>483</td>
</tr>
<tr>
<td></td>
<td>Convection</td>
<td>487</td>
</tr>
<tr>
<td></td>
<td>Radiation</td>
<td>491</td>
</tr>
<tr>
<td>15</td>
<td>Thermodynamics</td>
<td>505</td>
</tr>
<tr>
<td></td>
<td>The First Law of Thermodynamics</td>
<td>506</td>
</tr>
<tr>
<td></td>
<td>The First Law of Thermodynamics and Some Simple Processes</td>
<td>510</td>
</tr>
<tr>
<td></td>
<td>Introduction to the Second Law of Thermodynamics: Heat Engines and Their Efficiency</td>
<td>517</td>
</tr>
<tr>
<td></td>
<td>Carnot's Perfect Heat Engine: The Second Law of Thermodynamics Revisited</td>
<td>522</td>
</tr>
<tr>
<td></td>
<td>Applications of Thermodynamics: Heat Pumps and Refrigerators</td>
<td>526</td>
</tr>
<tr>
<td></td>
<td>Entropy and the Second Law of Thermodynamics: Disorder and the Unavailability of Energy</td>
<td>530</td>
</tr>
<tr>
<td></td>
<td>Statistical Interpretation of Entropy and the Second Law of Thermodynamics: The Underlying Explanation</td>
<td>536</td>
</tr>
<tr>
<td>16</td>
<td>Oscillatory Motion and Waves</td>
<td>549</td>
</tr>
<tr>
<td></td>
<td>Hooke's Law: Stress and Strain Revisited</td>
<td>550</td>
</tr>
<tr>
<td></td>
<td>Period and Frequency in Oscillations</td>
<td>554</td>
</tr>
<tr>
<td></td>
<td>Simple Harmonic Motion: A Special Periodic Motion</td>
<td>555</td>
</tr>
<tr>
<td></td>
<td>The Simple Pendulum</td>
<td>559</td>
</tr>
<tr>
<td></td>
<td>Energy and the Simple Harmonic Oscillator</td>
<td>561</td>
</tr>
<tr>
<td></td>
<td>Uniform Circular Motion and Simple Harmonic Motion</td>
<td>563</td>
</tr>
<tr>
<td></td>
<td>Damped Harmonic Motion</td>
<td>566</td>
</tr>
<tr>
<td></td>
<td>Forced Oscillations and Resonance</td>
<td>569</td>
</tr>
<tr>
<td></td>
<td>Waves</td>
<td>571</td>
</tr>
<tr>
<td></td>
<td>Superposition and Interference</td>
<td>573</td>
</tr>
<tr>
<td></td>
<td>Energy in Waves: Intensity</td>
<td>577</td>
</tr>
<tr>
<td>17</td>
<td>Physics of Hearing</td>
<td>589</td>
</tr>
<tr>
<td></td>
<td>Sound</td>
<td>590</td>
</tr>
<tr>
<td></td>
<td>Speed of Sound, Frequency, and Wavelength</td>
<td>591</td>
</tr>
<tr>
<td></td>
<td>Sound Intensity and Sound Level</td>
<td>594</td>
</tr>
<tr>
<td></td>
<td>Doppler Effect and Sonic Booms</td>
<td>598</td>
</tr>
<tr>
<td></td>
<td>Sound Interference and Resonance: Standing Waves in Air Columns</td>
<td>602</td>
</tr>
<tr>
<td></td>
<td>Hearing</td>
<td>608</td>
</tr>
<tr>
<td></td>
<td>Ultrasound</td>
<td>613</td>
</tr>
<tr>
<td>18</td>
<td>Electric Charge and Electric Field</td>
<td>627</td>
</tr>
<tr>
<td></td>
<td>Static Electricity and Charge: Conservation of Charge</td>
<td>628</td>
</tr>
<tr>
<td></td>
<td>Conductors and Insulators</td>
<td>632</td>
</tr>
<tr>
<td></td>
<td>Coulomb's Law</td>
<td>636</td>
</tr>
<tr>
<td></td>
<td>Electric Field: Concept of a Field Revisited</td>
<td>637</td>
</tr>
<tr>
<td></td>
<td>Electric Field Lines: Multiple Charges</td>
<td>639</td>
</tr>
<tr>
<td></td>
<td>Electric Forces in Biology</td>
<td>642</td>
</tr>
<tr>
<td></td>
<td>Conductors and Electric Fields in Static Equilibrium</td>
<td>643</td>
</tr>
<tr>
<td></td>
<td>Applications of Electrostatics</td>
<td>647</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>19</td>
<td>Electric Potential and Electric Field</td>
<td>663</td>
</tr>
<tr>
<td></td>
<td>Electric Potential Energy: Potential Difference</td>
<td>663</td>
</tr>
<tr>
<td></td>
<td>Electric Potential in a Uniform Electric Field</td>
<td>668</td>
</tr>
<tr>
<td></td>
<td>Electrical Potential Due to a Point Charge</td>
<td>671</td>
</tr>
<tr>
<td></td>
<td>Equipotential Lines</td>
<td>673</td>
</tr>
<tr>
<td></td>
<td>Capacitors and Dielectrics</td>
<td>675</td>
</tr>
<tr>
<td></td>
<td>Capacitors in Series and Parallel</td>
<td>681</td>
</tr>
<tr>
<td></td>
<td>Energy Stored in Capacitors</td>
<td>684</td>
</tr>
<tr>
<td>20</td>
<td>Electric Current, Resistance, and Ohm's Law</td>
<td>695</td>
</tr>
<tr>
<td></td>
<td>Current</td>
<td>696</td>
</tr>
<tr>
<td></td>
<td>Ohm's Law: Resistance and Simple Circuits</td>
<td>701</td>
</tr>
<tr>
<td></td>
<td>Resistance and Resistivity</td>
<td>702</td>
</tr>
<tr>
<td></td>
<td>Electric Power and Energy</td>
<td>707</td>
</tr>
<tr>
<td></td>
<td>Alternating Current versus Direct Current</td>
<td>710</td>
</tr>
<tr>
<td></td>
<td>Electric Hazards and the Human Body</td>
<td>714</td>
</tr>
<tr>
<td></td>
<td>Nerve Conduction—Electrocardiograms</td>
<td>717</td>
</tr>
<tr>
<td>21</td>
<td>Circuits, Bioelectricity, and DC Instruments</td>
<td>733</td>
</tr>
<tr>
<td></td>
<td>Resistors in Series and Parallel</td>
<td>733</td>
</tr>
<tr>
<td></td>
<td>Electromotive Force: Terminal Voltage</td>
<td>741</td>
</tr>
<tr>
<td></td>
<td>Kirchhoff's Rules</td>
<td>748</td>
</tr>
<tr>
<td></td>
<td>DC Voltmeters and Ammeters</td>
<td>752</td>
</tr>
<tr>
<td></td>
<td>Null Measurements</td>
<td>756</td>
</tr>
<tr>
<td></td>
<td>DC Circuits Containing Resistors and Capacitors</td>
<td>759</td>
</tr>
<tr>
<td>22</td>
<td>Magnets</td>
<td>773</td>
</tr>
<tr>
<td></td>
<td>Ferromagnets and Electromagnets</td>
<td>774</td>
</tr>
<tr>
<td></td>
<td>Magnetic Fields and Magnetic Field Lines</td>
<td>779</td>
</tr>
<tr>
<td></td>
<td>Magnetic Field Strength: Force on a Moving Charge in a Magnetic Field</td>
<td>780</td>
</tr>
<tr>
<td></td>
<td>Force on a Moving Charge in a Magnetic Field: Examples and Applications</td>
<td>781</td>
</tr>
<tr>
<td></td>
<td>The Hall Effect</td>
<td>785</td>
</tr>
<tr>
<td></td>
<td>Magnetic Force on a Current-Carrying Conductor</td>
<td>788</td>
</tr>
<tr>
<td></td>
<td>Torque on a Current Loop: Motors and Meters</td>
<td>790</td>
</tr>
<tr>
<td></td>
<td>Magnetic Fields Produced by Currents: Ampère's Law</td>
<td>792</td>
</tr>
<tr>
<td></td>
<td>Magnetic Force between Two Parallel Conductors</td>
<td>796</td>
</tr>
<tr>
<td></td>
<td>More Applications of Magnetism</td>
<td>797</td>
</tr>
<tr>
<td>23</td>
<td>Electromagnetic Induction, AC Circuits, and Electrical Technologies</td>
<td>811</td>
</tr>
<tr>
<td></td>
<td>Induced Emf and Magnetic Flux</td>
<td>812</td>
</tr>
<tr>
<td></td>
<td>Faraday's Law of Induction: Lenz's Law</td>
<td>814</td>
</tr>
<tr>
<td></td>
<td>Motional Emf</td>
<td>816</td>
</tr>
<tr>
<td></td>
<td>Eddy Currents and Magnetic Damping</td>
<td>819</td>
</tr>
<tr>
<td></td>
<td>Electric Generators</td>
<td>822</td>
</tr>
<tr>
<td></td>
<td>Back Emf</td>
<td>825</td>
</tr>
<tr>
<td></td>
<td>Transformers</td>
<td>825</td>
</tr>
<tr>
<td></td>
<td>Electrical Safety: Systems and Devices</td>
<td>829</td>
</tr>
<tr>
<td></td>
<td>Inductance</td>
<td>833</td>
</tr>
<tr>
<td></td>
<td>RL Circuits</td>
<td>836</td>
</tr>
<tr>
<td></td>
<td>Reactance, Inductive and Capacitive</td>
<td>838</td>
</tr>
<tr>
<td></td>
<td>RLC Series AC Circuits</td>
<td>841</td>
</tr>
<tr>
<td>24</td>
<td>Electromagnetic Waves</td>
<td>859</td>
</tr>
<tr>
<td></td>
<td>Maxwell's Equations: Electromagnetic Waves Predicted and Observed</td>
<td>860</td>
</tr>
<tr>
<td></td>
<td>Production of Electromagnetic Waves</td>
<td>862</td>
</tr>
<tr>
<td></td>
<td>The Electromagnetic Spectrum</td>
<td>864</td>
</tr>
<tr>
<td></td>
<td>Energy in Electromagnetic Waves</td>
<td>875</td>
</tr>
<tr>
<td>25</td>
<td>Geometric Optics</td>
<td>885</td>
</tr>
<tr>
<td></td>
<td>The Ray Aspect of Light</td>
<td>886</td>
</tr>
<tr>
<td></td>
<td>The Law of Reflection</td>
<td>887</td>
</tr>
<tr>
<td></td>
<td>The Law of Refraction</td>
<td>889</td>
</tr>
<tr>
<td></td>
<td>Total Internal Reflection</td>
<td>893</td>
</tr>
<tr>
<td></td>
<td>Dispersion: The Rainbow and Prisms</td>
<td>898</td>
</tr>
<tr>
<td></td>
<td>Image Formation by Lenses</td>
<td>902</td>
</tr>
<tr>
<td></td>
<td>Image Formation by Mirrors</td>
<td>913</td>
</tr>
<tr>
<td>26</td>
<td>Vision and Optical Instruments</td>
<td>927</td>
</tr>
<tr>
<td></td>
<td>Physics of the Eye</td>
<td>928</td>
</tr>
<tr>
<td></td>
<td>Vision Correction</td>
<td>931</td>
</tr>
<tr>
<td></td>
<td>Color and Color Vision</td>
<td>934</td>
</tr>
<tr>
<td></td>
<td>Microscopes</td>
<td>937</td>
</tr>
<tr>
<td></td>
<td>Telescopes</td>
<td>942</td>
</tr>
<tr>
<td></td>
<td>Aberrations</td>
<td>945</td>
</tr>
<tr>
<td>27</td>
<td>Wave Optics</td>
<td>953</td>
</tr>
<tr>
<td></td>
<td>The Wave Aspect of Light: Interference</td>
<td>954</td>
</tr>
<tr>
<td></td>
<td>Huygens's Principle: Diffraction</td>
<td>955</td>
</tr>
<tr>
<td></td>
<td>Young's Double Slit Experiment</td>
<td>957</td>
</tr>
<tr>
<td></td>
<td>Multiple Slit Diffraction</td>
<td>961</td>
</tr>
</tbody>
</table>
Textbook Equity* (textbookequity.org) provides high quality academic content and free access to any learner at any location, at the minimum possible cost. Education of everyone is the only true means to eliminate poverty and oppression. Creating barriers to education, whether intentional or not, by policy or price, leaves millions of geniuses undiscovered. We want them discovered. *Fearlessly copy, print, remix.®

About OpenStax College

OpenStax College is a non-profit organization committed to improving student access to quality learning materials. Our free textbooks are developed and peer-reviewed by educators to ensure they are readable, accurate, and meet the scope and sequence requirements of modern college courses. Unlike traditional textbooks, OpenStax College resources live online and are owned by the community of educators using them. Through our partnerships with companies and foundations committed to reducing costs for students, OpenStax College is working to improve access to higher education for all. OpenStax College is an initiative of Rice University and is made possible through the generous support of several philanthropic foundations.

About This Book

Welcome to College Physics, an OpenStax College resource created with several goals in mind: accessibility, affordability, customization, and student engagement—all while encouraging learners toward high levels of learning. Instructors and students alike will find that this textbook offers a strong foundation in introductory physics, with algebra as a prerequisite. It is available for free online and in low-cost print and e-book editions.

To broaden access and encourage community curation, College Physics is “open source” licensed under a Creative Commons Attribution (CC-BY) license. Everyone is invited to submit examples, emerging research, and other feedback to enhance and strengthen the material and keep it current and relevant for today's students. You can make suggestions by contacting us at info@openstaxcollege.org. You can find the status of the project, as well as alternate versions, corrections, etc., on the StaxDash at http://openstaxcollege.org.

To the Student

This book is written for you. It is based on the teaching and research experience of numerous physicists and influenced by a strong recollection of their own struggles as students. After reading this book, we hope you see that physics is visible everywhere. Applications range from driving a car to launching a rocket, from a skater whirling on ice to a neutron star spinning in space, and from taking your temperature to taking a chest X-ray.

To the Instructor

This text is intended for one-year introductory courses requiring algebra and some trigonometry, but no calculus. OpenStax College provides the essential supplemental resources at http://openstaxcollege.org; however, we have pared down the number of supplements to keep costs low. College Physics can be easily customized for your course using Connexions (http://cnx.org/content/col11406). Simply select the content most relevant to your curriculum and create a textbook that speaks directly to the needs of your class.

General Approach

College Physics is organized such that topics are introduced conceptually with a steady progression to precise definitions and analytical applications. The analytical aspect (problem solving) is tied back to the conceptual before moving on to another topic. Each introductory chapter, for example, opens with an engaging photograph relevant to the subject of the chapter and interesting applications that are easy for most students to visualize.

Organization, Level, and Content

There is considerable latitude on the part of the instructor regarding the use, organization, level, and content of this book. By choosing the types of problems assigned, the instructor can determine the level of sophistication required of the student.

Concepts and Calculations

The ability to calculate does not guarantee conceptual understanding. In order to unify conceptual, analytical, and calculation skills within the learning process, we have integrated Strategies and Discussions throughout the text.

Modern Perspective

The chapters on modern physics are more complete than many other texts on the market, with an entire chapter devoted to medical applications of nuclear physics and another to particle physics. The final chapter of the text, “Frontiers of Physics,” is devoted to the most exciting endeavors in physics. It ends with a module titled “Some Questions We Know to Ask.”

Supplements

(Not necessarily endorsed by Textbook Equity)


Features of OpenStax College Physics

The following briefly describes the special features of this text.
Modularity

This textbook is organized on Connexions (http://cnx.org) as a collection of modules that can be rearranged and modified to suit the needs of a particular professor or class. That being said, modules often contain references to content in other modules, as most topics in physics cannot be discussed in isolation.

Learning Objectives

Every module begins with a set of learning objectives. These objectives are designed to guide the instructor in deciding what content to include or assign, and to guide the student with respect to what he or she can expect to learn. After completing the module and end-of-module exercises, students should be able to demonstrate mastery of the learning objectives.

Call-Outs

Key definitions, concepts, and equations are called out with a special design treatment. Call-outs are designed to catch readers’ attention, to make it clear that a specific term, concept, or equation is particularly important, and to provide easy reference for a student reviewing content.

Key Terms

Key terms are in bold and are followed by a definition in context. Definitions of key terms are also listed in the Glossary, which appears at the end of the module.

Worked Examples

Worked examples have four distinct parts to promote both analytical and conceptual skills. Worked examples are introduced in words, always using some application that should be of interest. This is followed by a Strategy section that emphasizes the concepts involved and how solving the problem relates to those concepts. This is followed by the mathematical Solution and Discussion.

Many worked examples contain multiple-part problems to help the students learn how to approach normal situations, in which problems tend to have multiple parts. Finally, worked examples employ the techniques of the problem-solving strategies so that students can see how those strategies succeed in practice as well as in theory.

Problem-Solving Strategies

Problem-solving strategies are first presented in a special section and subsequently appear at crucial points in the text where students can benefit most from them. Problem-solving strategies have a logical structure that is reinforced in the worked examples and supported in certain places by line drawings that illustrate various steps.

Misconception Alerts

Students come to physics with preconceptions from everyday experiences and from previous courses. Some of these preconceptions are misconceptions, and many are very common among students and the general public. Some are inadvertently picked up through misunderstandings of lectures and texts. The Misconception Alerts feature is designed to point these out and correct them explicitly.

Take-Home Investigations

Take Home Investigations provide the opportunity for students to apply or explore what they have learned with a hands-on activity.

Things Great and Small

In these special topic essays, macroscopic phenomena (such as air pressure) are explained with submicroscopic phenomena (such as atoms bouncing off walls). These essays support the modern perspective by describing aspects of modern physics before they are formally treated in later chapters. Connections are also made between apparently disparate phenomena.

Simulations

Where applicable, students are directed to the interactive PHeT physics simulations developed by the University of Colorado (http://phet.colorado.edu). There they can further explore the physics concepts they have learned about in the module.

Summary

Module summaries are thorough and functional and present all important definitions and equations. Students are able to find the definitions of all terms and symbols as well as their physical relationships. The structure of the summary makes plain the fundamental principles of the module or collection and serves as a useful study guide.

Glossary

At the end of every module or chapter is a glossary containing definitions of all of the key terms in the module or chapter.

End-of-Module Problems

At the end of every chapter is a set of Conceptual Questions and/or skills-based Problems & Exercises. Conceptual Questions challenge students’ ability to explain what they have learned conceptually, independent of the mathematical details. Problems & Exercises challenge students to apply both concepts and skills to solve mathematical physics problems. Online, every other problem includes an answer that students can reveal immediately by clicking on a “Show Solution” button. Fully worked solutions to select problems are available in the Student Solutions Manual and the Teacher Solutions Manual.

In addition to traditional skills-based problems, there are three special types of end-of-module problems: Integrated Concept Problems, Unreasonable Results Problems, and Construct Your Own Problems. All of these problems are indicated with a subtitle preceding the problem.
Integrated Concept Problems

In Unreasonable Results Problems, students are challenged not only to apply concepts and skills to solve a problem, but also to analyze the answer with respect to how likely or realistic it really is. These problems contain a premise that produces an unreasonable answer and are designed to further emphasize that properly applied physics must describe nature accurately and is not simply the process of solving equations.

Unreasonable Results

In Unreasonable Results Problems, students are challenged not only to apply concepts and skills to solve a problem, but also to analyze the answer with respect to how likely or realistic it really is. These problems contain a premise that produces an unreasonable answer and are designed to further emphasize that properly applied physics must describe nature accurately and is not simply the process of solving equations.

Construct Your Own Problem

These problems require students to construct the details of a problem, justify their starting assumptions, show specific steps in the problem’s solution, and finally discuss the meaning of the result. These types of problems relate well to both conceptual and analytical aspects of physics, emphasizing that physics must describe nature. Often they involve an integration of topics from more than one chapter. Unlike other problems, solutions are not provided since there is no single correct answer. Instructors should feel free to direct students regarding the level and scope of their considerations. Whether the problem is solved and described correctly will depend on initial assumptions.

Appendices

Appendix A: Atomic Masses
Appendix B: Selected Radioactive Isotopes
Appendix C: Useful Information
Appendix D: Glossary of Key Symbols and Notation

Acknowledgements

This text is based on the work completed by Dr. Paul Peter Urone in collaboration with Roger Hinrichs, Kim Dirks, and Manjula Sharma. We would like to thank the authors as well as the numerous professors (a partial list follows) who have contributed their time and energy to review and provide feedback on the manuscript. Their input has been critical in maintaining the pedagogical integrity and accuracy of the text.

Senior Contributing Authors

Dr. Paul Peter Urone
Dr. Roger Hinrichs, State University of New York, College at Oswego

Contributing Authors

Dr. Kim Dirks, University of Auckland, New Zealand
Dr. Manjula Sharma, University of Sydney, Australia

Expert Reviewers

Erik Christensen, P.E, South Florida Community College
Dr. Eric Kincanon, Gonzaga University
Dr. Douglas Ingram, Texas Christian University
Lee H. LaRue, Paris Junior College
Dr. Marc Sher, College of William and Mary
Dr. Ulrich Zurcher, Cleveland State University
Dr. Matthew Adams, Crafton Hills College, San Bernardino Community College District
Dr. Chuck Pearson, Virginia Intermont College

Our Partners

(Not necessarily endorsed by Textbook Equity)

WebAssign

WebAssign is an independent online homework and assessment system that has been available commercially since 1998. WebAssign has recently begun to support the Open Education Resource community by creating a high quality online homework solution for selected open-source textbooks, available at an affordable price to students. These question collections include randomized values and variables, immediate feedback, links to the open-source textbook, and a variety of text-specific resources and tools; as well as the same level of rigorous coding and accuracy-checking as any commercially available online homework solution supporting traditionally available textbooks.

Sapling Learning

Sapling Learning provides the most effective interactive homework and instruction that improve student learning outcomes for the problem-solving disciplines. They offer an enjoyable teaching and effective learning experience that is distinctive in three important ways:

- Ease of Use: Sapling Learning’s easy to use interface keeps students engaged in problem-solving, not struggling with the software.
- Targeted Instructional Content: Sapling Learning increases student engagement and comprehension by delivering immediate feedback and targeted instructional content.
- Unsurpassed Service and Support: Sapling Learning makes teaching more enjoyable by providing a dedicated Masters or PhD level colleague to service instructors’ unique needs throughout the course, including content customization.
18. Your friend says that more patterns and colors can be seen on the wings of birds if viewed in ultraviolet light. Would you agree with your friend? Explain your answer.

19. The rate at which information can be transmitted on an electromagnetic wave is proportional to the frequency of the wave. Is this consistent with the fact that laser telephone transmission at visible frequencies carries far more conversations per optical fiber than conventional electronic transmission in a wire? What is the implication for ELF radio communication with submarines?

20. Give an example of energy carried by an electromagnetic wave.

21. In an MRI scan, a higher magnetic field requires higher frequency radio waves to resonate with the nuclear type whose density and location is being imaged. What effect does going to a larger magnetic field have on the most efficient antenna to broadcast those radio waves? Does it favor a smaller or larger antenna?

22. Laser vision correction often uses an excimer laser that produces 193-nm electromagnetic radiation. This wavelength is extremely strongly absorbed by the cornea and ablates it in a manner that reshapes the cornea to correct vision defects. Explain how the strong absorption helps concentrate the energy in a thin layer and thus give greater accuracy in shaping the cornea. Also explain how this strong absorption limits damage to the lens and retina of the eye.
Problems & Exercises

24.1 Maxwell’s Equations: Electromagnetic Waves Predicted and Observed

1. Verify that the correct value for the speed of light \( c \) is obtained when numerical values for the permeability and permittivity of free space (\( \mu_0 \) and \( \epsilon_0 \)) are entered into the equation \( c = \sqrt{\frac{1}{\mu_0 \epsilon_0}} \).

2. Show that, when SI units for \( \mu_0 \) and \( \epsilon_0 \) are entered, the units given by the right-hand side of the equation in the problem above are m/s.

24.2 Production of Electromagnetic Waves

3. What is the maximum electric field strength in an electromagnetic wave that has a maximum magnetic field strength of \( 5.00 \times 10^{-4} \text{ T} \)?

4. The maximum field strength of an electromagnetic field is \( 5 \times 10^{-6} \text{ T} \). Calculate the maximum electric field strength if the wave is traveling in a medium in which the speed of the wave is \( 0.75c \).

5. Verify the units obtained for magnetic field strength \( B \) in Example 24.1 (using the equation \( B = \frac{E}{c^2} \)) are in fact teslas (T).

24.3 The Electromagnetic Spectrum

6. (a) Two microwave frequencies are authorized for use in microwave ovens: 900 and 2560 MHz. Calculate the wavelength of each. (b) Which frequency would produce smaller hot spots in foods due to interference effects?

7. (a) Calculate the range of frequencies for AM radio given its frequency range is 540 to 1600 kHz. (b) Do the same for the FM frequency range of 88.0 to 108 MHz.

8. A radio station utilizes frequencies between commercial AM and FM. What is the frequency of a 11.12-m-wavelength channel?

9. Find the frequency range of visible light, given that it encompasses wavelengths from 380 to 760 nm.

10. Combing your hair leads to excess electrons on the comb. How fast would you have to move the comb up and down to produce red light?

11. Electromagnetic radiation having a \( 15.0 \mu \text{m} \) wavelength is classified as infrared radiation. What is its frequency?

12. Approximately what is the smallest detail observable with a microscope that uses ultraviolet light of frequency \( 1.20 \times 10^{15} \text{ Hz} \)?

13. A radar used to detect the presence of aircraft receives a pulse that has reflected off an object \( 6 \times 10^{-5} \text{ s} \) after it was transmitted. What is the distance from the radar station to the reflecting object?

14. Some radar systems detect the size and shape of objects such as aircraft and geological terrain. Approximately what is the smallest observable detail utilizing 500-MHz radar?

15. Determine the amount of time it takes for X-rays of frequency \( 3 \times 10^{18} \text{ Hz} \) to travel (a) 1 mm and (b) 1 cm.

16. If you wish to detect details of the size of atoms (about \( 1 \times 10^{-10} \text{ m} \)) with electromagnetic radiation, it must have a wavelength of about this size. (a) What is its frequency? (b) What type of electromagnetic radiation might this be?

17. If the Sun suddenly turned off, we would not know it until its light stopped coming. How long would that be, given that the Sun is \( 1.50 \times 10^{11} \text{ m} \) away?

18. Distances in space are often quoted in units of light years, the distance light travels in one year. (a) How many meters is a light year?

19. A certain 50.0-Hz AC power line radiates an electromagnetic wave having a maximum electric field strength of \( 13.0 \text{ kV/m} \). (a) What is the wavelength of this very low frequency electromagnetic wave? (b) What is its maximum magnetic field strength?

20. During normal beating, the heart creates a maximum 4.00-mV potential across 0.300 m of a person’s chest, creating a 1.00-Hz electromagnetic wave. (a) What is the maximum electric field strength created? (b) What is the corresponding maximum magnetic field strength in the electromagnetic wave? (c) What is the wavelength of the electromagnetic wave?

21. (a) The ideal size (most efficient) for a broadcast antenna with one end on the ground is one-fourth the wavelength \( (\lambda/4) \) of the electromagnetic radiation being sent out. If a new radio station has such an antenna that is 50.0 m high, what frequency does it broadcast most efficiently? Is this in the AM or FM band? (b) Discuss the analogy of the fundamental resonant mode of an air column closed at one end to the resonance of currents on an antenna that is one-fourth their wavelength.

22. (a) What is the wavelength of 100-MHz radio waves used in an MRI unit? (b) If the frequencies are swept over a \( \pm 1.00 \text{ kHz} \) range centered on 100 MHz, what is the range of wavelengths broadcast?

23. (a) What is the frequency of the 193-nm ultraviolet radiation used in laser eye surgery? (b) Assuming the accuracy with which this EM radiation can ablate the cornea is directly proportional to wavelength, how much more accurate can this UV be than the shortest visible wavelength of light?

24. TV-reception antennas for VHF are constructed with cross wires supported at their centers, as shown in Figure 24.27. The ideal length for the cross wires is one-half the wavelength to be received, with the more expensive antennas having one for each channel. Suppose you measure the lengths of the wires for particular channels and find them to be 1.94 and 0.753 m long, respectively. What are the frequencies for these channels?

25. Conversations with astronauts on lunar walks had an echo that was used to estimate the distance to the Moon. The sound spoken by the person on Earth was transformed into a radio signal sent to the Moon, and transformed back into sound on a speaker inside the astronaut’s space suit. This sound was picked up by the microphone in the space suit (intended for the astronaut’s voice) and sent back to Earth as a radio echo of sorts. If the round-trip time was 2.60 s, what was the approximate distance to the Moon, neglecting any delays in the electronics?

26. Lunar astronauts placed a reflector on the Moon’s surface, off which a laser beam is periodically reflected. The distance to the Moon is calculated from the round-trip time. (a) To what accuracy in meters can the distance to the Moon be determined, if this time can be measured to
0.100 ns? (b) What percent accuracy is this, given the average distance to the Moon is 3.84×10^8 m?

27. Radar is used to determine distances to various objects by measuring the round-trip time for an echo from the object. (a) How far away is the planet Venus if the echo time is 1.00 s? (b) What is the echo time for a car 75.0 m from a Highway Police radar unit? (c) How accurately (in nanoseconds) must you be able to measure the echo time to an airplane 12.0 km away to determine its distance within 10.0 m?

28. Integrated Concepts
(a) Calculate the ratio of the highest to lowest frequencies of electromagnetic waves the eye can see, given the wavelength range of visible light is from 380 to 760 nm. (b) Compare this with the ratio of highest to lowest frequencies the ear can hear.

29. Integrated Concepts
(a) Calculate the rate in watts at which heat transfer through radiation occurs (almost entirely in the infrared) from 1.0 m^2 of the Earth’s surface at night. Assume the emissivity is 0.90, the temperature of the Earth is 15°C, and that of outer space is 2.7 K. (b) Compare the intensity of this radiation with that coming to the Earth from the Sun during the day, which averages about 800 W/m^2, only half of which is absorbed. (c) What is the maximum magnetic field strength in the outgoing radiation, assuming it is a continuous wave?

24.4 Energy in Electromagnetic Waves
30. What is the intensity of an electromagnetic wave with a peak electric field strength of 125 V/m?

31. Find the intensity of an electromagnetic wave having a peak magnetic field strength of 4.00×10^{-9} T.

32. Assume the helium-neon lasers commonly used in student physics laboratories have power outputs of 0.500 mW. (a) If such a laser beam is projected onto a circular spot 1.00 mm in diameter, what is its intensity? (b) Find the peak magnetic field strength. (c) Find the peak electric field strength.

33. An AM radio transmitter broadcasts 50.0 kW of power uniformly in all directions. (a) Assuming all of the radio waves that strike the ground are completely absorbed, and that there is no absorption by the atmosphere or other objects, what is the intensity 30.0 km away? (Hint: Half the power will be spread over the area of a hemisphere.) (b) What is the maximum electric field strength at this distance?

34. Suppose the maximum safe intensity of microwaves for human exposure is taken to be 1.00 W/m^2. (a) If a radar unit leaks 10.0 W of microwaves (other than those sent by its antenna) uniformly in all directions, how far away must you be to be exposed to an intensity considered to be safe? Assume that the power spreads uniformly over the area of a sphere with no complications from absorption or reflection. (b) What is the maximum electric field strength at the safe intensity? (Note that early radar units leaked more than modern ones do. This caused identifiable health problems, such as cataracts, for people who worked near them.)

35. A 2.50-m-diameter university communications satellite dish receives TV signals that have a maximum electric field strength (for one channel) of 7.50 μV/m. (See Figure 24.28.) (a) What is the intensity of this wave? (b) What is the power received by the antenna? (c) If the orbiting satellite broadcasts uniformly over an area of 1.50×10^{13} m^2 (a large fraction of North America), how much power does it radiate?

Figure 24.28 Satellite dishes receive TV signals sent from orbit. Although the signals are quite weak, the receiver can detect them by being tuned to resonate at their frequency.

36. Lasers can be constructed that produce an extremely high intensity electromagnetic wave for a brief time—called pulsed lasers. They are used to ignite nuclear fusion, for example. Such a laser may produce an electromagnetic wave with a maximum electric field strength of 1.00×10^{11} V/m for a time of 1.00 ns. (a) What is the maximum magnetic field strength in the wave? (b) What is the intensity of the beam? (c) What energy does it deliver on a 1.00-mm^2 area?

37. Show that for a continuous sinusoidal electromagnetic wave, the peak intensity is twice the average intensity (I_0 = 2I_{ave}), using either the fact that E_0 = √2E_{rms}, or B_0 = √2B_{rms}, where rms means average (actually root mean square, a type of average).

38. Suppose a source of electromagnetic waves radiates uniformly in all directions in empty space where there are no absorption or interference effects. (a) Show that the intensity is inversely proportional to r^2, the distance from the source squared. (b) Show that the magnitudes of the electric and magnetic fields are inversely proportional to r.

39. Integrated Concepts
An LC circuit with a 5.00-pF capacitor oscillates in such a manner as to radiate at a wavelength of 3.30 m. (a) What is the resonant frequency? (b) What inductance is in series with the capacitor?

40. Integrated Concepts
What capacitance is needed in series with an 800 – μH inductor to form a circuit that radiates a wavelength of 196 m?

41. Integrated Concepts
Police radar determines the speed of motor vehicles using the same Doppler-shift technique employed for ultrasound in medical diagnostics. Beats are produced by mixing the double Doppler-shifted echo with the original frequency. If 1.50×10^{10}-Hz microwaves are used and a beat frequency of 150 Hz is produced, what is the speed of the vehicle? (Assume the same Doppler-shift formulas are valid with the speed of sound replaced by the speed of light.)

42. Integrated Concepts
Assume the mostly infrared radiation from a heat lamp acts like a continuous wave with wavelength 1.50 μm. (a) If the lamp’s 200-W output is focused on a person’s shoulder, over a circular area 25.0 cm in diameter, what is the intensity in W/m^2? (b) What is the peak electric field strength? (c) Find the peak magnetic field strength. (d) How long will it take to increase the temperature of the 4.00-kg shoulder by 2.00°C?
assuming no other heat transfer and given that its specific heat is $3.47 \times 10^3$ J/kg \cdot °C ?

43. Integrated Concepts
On its highest power setting, a microwave oven increases the temperature of 0.400 kg of spaghetti by 45.0ºC in 120 s. (a) What was the rate of power absorption by the spaghetti, given that its specific heat is $3.76 \times 10^3$ J/kg \cdot °C ? (b) Find the average intensity of the microwaves, given that they are absorbed over a circular area 20.0 cm in diameter. (c) What is the peak electric field strength of the microwave? (d) What is its peak magnetic field strength?

44. Integrated Concepts
Electromagnetic radiation from a 5.00-mW laser is concentrated on a 1.00-mm$^2$ area. (a) What is the intensity in W/m$^2$? (b) Suppose a 2.00-nC static charge is in the beam. What is the maximum electric force it experiences? (c) If the static charge moves at 400 m/s, what maximum magnetic force can it feel?

45. Integrated Concepts
A 200-turn flat coil of wire 30.0 cm in diameter acts as an antenna for FM radio at a frequency of 100 MHz. The magnetic field of the incoming electromagnetic wave is perpendicular to the coil and has a maximum strength of $1.00 \times 10^{-12}$ T. (a) What power is incident on the coil? (b) What average emf is induced in the coil over one-fourth of a cycle? (c) If the radio receiver has an inductance of 2.50 µH, what capacitance must it have to resonate at 100 MHz?

46. Integrated Concepts
If electric and magnetic field strengths vary sinusoidally in time, being zero at $t = 0$, then $E = E_0 \sin 2\pi ft$ and $B = B_0 \sin 2\pi ft$. Let $f = 1.00$ GHz here. (a) When are the field strengths first zero? (b) When do they reach their most negative value? (c) How much time is needed for them to complete one cycle?

47. Unreasonable Results
A researcher measures the wavelength of a 1.20-GHz electromagnetic wave to be 0.500 m. (a) Calculate the speed at which this wave propagates. (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

48. Unreasonable Results
The peak magnetic field strength in a residential microwave oven is $9.20 \times 10^{-5}$ T. (a) What is the intensity of the microwave? (b) What is unreasonable about this result? (c) What is wrong about the premise?

49. Unreasonable Results
An LC circuit containing a 2.00-H inductor oscillates at such a frequency that it radiates at a 1.00-m wavelength. (a) What is the capacitance of the circuit? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

50. Unreasonable Results
An LC circuit containing a 1.00-pF capacitor oscillates at such a frequency that it radiates at a 300-nm wavelength. (a) What is the inductance of the circuit? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

51. Create Your Own Problem
Consider electromagnetic fields produced by high voltage power lines. Construct a problem in which you calculate the intensity of this electromagnetic radiation in W/m$^2$ based on the measured magnetic field strength of the radiation in a home near the power lines. Assume these magnetic field strengths are known to average less than a µT. The intensity is small enough that it is difficult to imagine mechanisms for biological damage due to it. Discuss how much energy may be radiating from a section of power line several hundred meters long and compare this to the power likely to be carried by the lines. An idea of how much power this is can be obtained by calculating the approximate current responsible for µT fields at distances of tens of meters.
Learning Objectives

25.1. The Ray Aspect of Light
• List the ways by which light travels from a source to another location.

25.2. The Law of Reflection
• Explain reflection of light from polished and rough surfaces.

25.3. The Law of Refraction
• Determine the index of refraction, given the speed of light in a medium.

25.4. Total Internal Reflection
• Explain the phenomenon of total internal reflection.
• Describe the workings and uses of fiber optics.
• Analyze the reason for the sparkle of diamonds.

25.5. Dispersion: The Rainbow and Prisms
• Explain the phenomenon of dispersion and discuss its advantages and disadvantages.

25.6. Image Formation by Lenses
• List the rules for ray tracking for thin lenses.
• Illustrate the formation of images using the technique of ray tracking.
• Determine power of a lens given the focal length.

25.7. Image Formation by Mirrors
• Illustrate image formation in a flat mirror.
• Explain with ray diagrams the formation of an image using spherical mirrors.
• Determine focal length and magnification given radius of curvature, distance of object and image.

Introduction to Geometric Optics

Geometric Optics

Light from this page or screen is formed into an image by the lens of your eye, much as the lens of the camera that made this photograph. Mirrors, like lenses, can also form images that in turn are captured by your eye.
Our lives are filled with light. Through vision, the most valued of our senses, light can evoke spiritual emotions, such as when we view a magnificent sunset or glimpse a rainbow breaking through the clouds. Light can also simply amuse us in a theater, or warn us to stop at an intersection. It has innumerable uses beyond vision. Light can carry telephone signals through glass fibers or cook a meal in a solar oven. Life itself could not exist without light's energy. From photosynthesis in plants to the sun warming a cold-blooded animal, its supply of energy is vital.

Figure 25.2 Double Rainbow over the bay of Pocitos in Montevideo, Uruguay. (credit: Madrax, Wikimedia Commons)

We already know that visible light is the type of electromagnetic waves to which our eyes respond. That knowledge still leaves many questions regarding the nature of light and vision. What is color, and how do our eyes detect it? Why do diamonds sparkle? How does light travel? How do lenses and mirrors form images? These are but a few of the questions that are answered by the study of optics. Optics is the branch of physics that deals with the behavior of visible light and other electromagnetic waves. In particular, optics is concerned with the generation and propagation of light and its interaction with matter. What we have already learned about the generation of light in our study of heat transfer by radiation will be expanded upon in later topics, especially those on atomic physics. Now, we will concentrate on the propagation of light and its interaction with matter.

It is convenient to divide optics into two major parts based on the size of objects that light encounters. When light interacts with an object that is several times as large as the light's wavelength, its observable behavior is like that of a ray; it does not prominently display its wave characteristics. We call this part of optics “geometric optics.” This chapter will concentrate on such situations. When light interacts with smaller objects, it has very prominent wave characteristics, such as constructive and destructive interference. Wave Optics will concentrate on such situations.

25.1 The Ray Aspect of Light

There are three ways in which light can travel from a source to another location. (See Figure 25.3.) It can come directly from the source through empty space, such as from the Sun to Earth. Or light can travel through various media, such as air and glass, to the person. Light can also arrive after being reflected, such as by a mirror. In all of these cases, light is modeled as traveling in straight lines called rays. Light may change direction when it encounters objects (such as a mirror) or in passing from one material to another (such as in passing from air to glass), but it then continues in a straight line or as a ray. The word ray comes from mathematics and here means a straight line that originates at some point. It is acceptable to visualize light rays as laser rays (or even science fiction depictions of ray guns).

Ray

The word “ray” comes from mathematics and here means a straight line that originates at some point.

Figure 25.3 Three methods for light to travel from a source to another location. (a) Light reaches the upper atmosphere of Earth traveling through empty space directly from the source. (b) Light can reach a person in one of two ways. It can travel through media like air and glass. It can also reflect from an object like a mirror. In the situations shown here, light interacts with objects large enough that it travels in straight lines, like a ray.

Experiments, as well as our own experiences, show that when light interacts with objects several times as large as its wavelength, it travels in straight lines and acts like a ray. Its wave characteristics are not pronounced in such situations. Since the wavelength of light is less than a micron (a thousandth of a millimeter), it acts like a ray in the many common situations in which it encounters objects larger than a micron. For example, when light encounters anything we can observe with unaided eyes, such as a mirror, it acts like a ray, with only subtle wave characteristics. We will concentrate on the ray characteristics in this chapter.

Since light moves in straight lines, changing directions when it interacts with materials, it is described by geometry and simple trigonometry. This part of optics, where the ray aspect of light dominates, is therefore called geometric optics. There are two laws that govern how light changes direction...
when it interacts with matter. These are the law of reflection, for situations in which light bounces off matter, and the law of refraction, for situations in which light passes through matter.

### Geometric Optics

The part of optics dealing with the ray aspect of light is called geometric optics.

#### 25.2 The Law of Reflection

Whenever we look into a mirror, or squint at sunlight glinting from a lake, we are seeing a reflection. When you look at this page, too, you are seeing light reflected from it. Large telescopes use reflection to form an image of stars and other astronomical objects.

The law of reflection is illustrated in Figure 25.4, which also shows how the angles are measured relative to the perpendicular to the surface at the point where the light ray strikes. We expect to see reflections from smooth surfaces, but Figure 25.5 illustrates how a rough surface reflects light. Since the light strikes different parts of the surface at different angles, it is reflected in many different directions, or diffused. Diffused light is what allows us to see a sheet of paper from any angle, as illustrated in Figure 25.6. Many objects, such as people, clothing, leaves, and walls, have rough surfaces and can be seen from all sides. A mirror, on the other hand, has a smooth surface (compared with the wavelength of light) and reflects light at specific angles, as illustrated in Figure 25.7. When the moon reflects from a lake, as shown in Figure 25.8, a combination of these effects takes place.

#### Figure 25.4

The law of reflection states that the angle of reflection equals the angle of incidence—\( \theta_r = \theta_i \). The angles are measured relative to the perpendicular to the surface at the point where the ray strikes the surface.

#### Figure 25.5

Light is diffused when it reflects from a rough surface. Here many parallel rays are incident, but they are reflected at many different angles since the surface is rough.

#### Figure 25.6

When a sheet of paper is illuminated with many parallel incident rays, it can be seen at many different angles, because its surface is rough and diffuses the light.
The law of reflection is very simple: The angle of reflection equals the angle of incidence.

When we see ourselves in a mirror, it appears that our image is actually behind the mirror. This is illustrated in Figure 25.9. We see the light coming from a direction determined by the law of reflection. The angles are such that our image is exactly the same distance behind the mirror as we stand away from the mirror. If the mirror is on the wall of a room, the images in it are all behind the mirror, which can make the room seem bigger. Although these mirror images make objects appear to be where they cannot be (like behind a solid wall), the images are not figments of our imagination. Mirror images can be photographed and videotaped by instruments and look just as they do with our eyes (optical instruments themselves). The precise manner in which images are formed by mirrors and lenses will be treated in later sections of this chapter.

Figure 25.9 Our image in a mirror is behind the mirror. The two rays shown are those that strike the mirror at just the correct angles to be reflected into the eyes of the person. The image appears to be in the direction the rays are coming from when they enter the eyes.

**Take-Home Experiment: Law of Reflection**

Take a piece of paper and shine a flashlight at an angle at the paper, as shown in Figure 25.6. Now shine the flashlight at a mirror at an angle. Do your observations confirm the predictions in Figure 25.6 and Figure 25.7? Shine the flashlight on various surfaces and determine whether the reflected light is diffuse or not. You can choose a shiny metallic lid of a pot or your skin. Using the mirror and flashlight, can you confirm the
25.3 The Law of Refraction

It is easy to notice some odd things when looking into a fish tank. For example, you may see the same fish appearing to be in two different places. (See Figure 25.10.) This is because light coming from the fish to us changes direction when it leaves the tank, and in this case, it can travel two different paths to get to our eyes. The changing of a light ray’s direction (loosely called bending) when it passes through variations in matter is called refraction. Refraction is responsible for a tremendous range of optical phenomena, from the action of lenses to voice transmission through optical fibers.

Refraction

The changing of a light ray’s direction (loosely called bending) when it passes through variations in matter is called refraction.

Speed of Light

The speed of light $c$ not only affects refraction, it is one of the central concepts of Einstein’s theory of relativity. As the accuracy of the measurements of the speed of light were improved, $c$ was found not to depend on the velocity of the source or the observer. However, the speed of light does vary in a precise manner with the material it traverses. These facts have far-reaching implications, as we will see in Special Relativity. It makes connections between space and time and alters our expectations that all observers measure the same time for the same event, for example. The speed of light is so important that its value in a vacuum is one of the most fundamental constants in nature as well as being one of the four fundamental SI units.

Figure 25.10 Looking at the fish tank as shown, we can see the same fish in two different locations, because light changes directions when it passes from water to air. In this case, the light can reach the observer by two different paths, and so the fish seems to be in two different places. This bending of light is called refraction and is responsible for many optical phenomena.

Why does light change direction when passing from one material (medium) to another? It is because light changes speed when going from one material to another. So before we study the law of refraction, it is useful to discuss the speed of light and how it varies in different media.

The Speed of Light

Early attempts to measure the speed of light, such as those made by Galileo, determined that light moved extremely fast, perhaps instantaneously. The first real evidence that light traveled at a finite speed came from the Danish astronomer Ole Roemer in the late 17th century. Roemer had noted that the average orbital period of one of Jupiter’s moons, as measured from Earth, varied depending on whether Earth was moving toward or away from Jupiter. He correctly concluded that the apparent change in period was due to the change in distance between Earth and Jupiter and the time it took light to travel this distance. From his 1676 data, a value of the speed of light was calculated to be $2.26 \times 10^8$ m/s (only 25% different from today’s accepted value). In more recent times, physicists have measured the speed of light in numerous ways and with increasing accuracy. One particularly direct method, used in 1887 by the American physicist Albert Michelson (1852–1931), is illustrated in Figure 25.11. Light reflected from a rotating set of mirrors was reflected from a stationary mirror 35 km away and returned to the rotating mirrors. The time for the light to travel can be determined by how fast the mirrors must rotate for the light to be returned to the observer’s eye.
Figure 25.11 A schematic of early apparatus used by Michelson and others to determine the speed of light. As the mirrors rotate, the reflected ray is only briefly directed at the stationary mirror. The returning ray will be reflected into the observer’s eye only if the next mirror has rotated into the correct position just as the ray returns. By measuring the correct rotation rate, the time for the round trip can be measured and the speed of light calculated. Michelson’s calculated value of the speed of light was only 0.04% different from the value used today.

The speed of light is now known to great precision. In fact, the speed of light in a vacuum \( c \) is so important that it is accepted as one of the basic physical quantities and has the fixed value

\[
c = 2.9972458 \times 10^8 \text{ m/s} \approx 3.00 \times 10^8 \text{ m/s},
\]

where the approximate value of \( 3.00 \times 10^8 \text{ m/s} \) is used whenever three-digit accuracy is sufficient. The speed of light through matter is less than it is in a vacuum, because light interacts with atoms in a material. The speed of light depends strongly on the type of material, since its interaction with different atoms, crystal lattices, and other substructures varies. We define the index of refraction \( n \) of a material to be

\[
n = \frac{c}{v},
\]

where \( v \) is the observed speed of light in the material. Since the speed of light is always less than \( c \) in matter and equals \( c \) only in a vacuum, the index of refraction is always greater than or equal to one.

<table>
<thead>
<tr>
<th>Value of the Speed of Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c = 2.9972458 \times 10^8 \text{ m/s} \approx 3.00 \times 10^8 \text{ m/s} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Index of Refraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = \frac{c}{v} )</td>
</tr>
</tbody>
</table>

That is, \( n \geq 1 \). Table 25.1 gives the indices of refraction for some representative substances. The values are listed for a particular wavelength of light, because they vary slightly with wavelength. (This can have important effects, such as colors produced by a prism.) Note that for gases, \( n \) is close to 1.0. This seems reasonable, since atoms in gases are widely separated and light travels at \( c \) in the vacuum between atoms. It is common to take \( n = 1 \) for gases unless great precision is needed. Although the speed of light \( v \) in a medium varies considerably from its value \( c \) in a vacuum, it is still a large speed.
Table 25.1 Index of Refraction in Various Media

<table>
<thead>
<tr>
<th>Medium</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gases at 0ºC, 1 atm</strong></td>
<td></td>
</tr>
<tr>
<td>Air</td>
<td>1.000293</td>
</tr>
<tr>
<td>Carbon dioxide</td>
<td>1.00045</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>1.000139</td>
</tr>
<tr>
<td>Oxygen</td>
<td>1.000271</td>
</tr>
<tr>
<td>** Liquids at 20ºC    **</td>
<td></td>
</tr>
<tr>
<td>Benzene</td>
<td>1.501</td>
</tr>
<tr>
<td>Carbon disulfide</td>
<td>1.628</td>
</tr>
<tr>
<td>Carbon tetrachloride</td>
<td>1.461</td>
</tr>
<tr>
<td>Ethanol</td>
<td>1.361</td>
</tr>
<tr>
<td>Glycerine</td>
<td>1.473</td>
</tr>
<tr>
<td>Water, fresh</td>
<td>1.333</td>
</tr>
<tr>
<td>** Solids at 20ºC     **</td>
<td></td>
</tr>
<tr>
<td>Diamond</td>
<td>2.419</td>
</tr>
<tr>
<td>Fluorite</td>
<td>1.434</td>
</tr>
<tr>
<td>Glass, crown</td>
<td>1.52</td>
</tr>
<tr>
<td>Glass, flint</td>
<td>1.66</td>
</tr>
<tr>
<td>Ice at 20ºC</td>
<td>1.309</td>
</tr>
<tr>
<td>Polystyrene</td>
<td>1.49</td>
</tr>
<tr>
<td>Plexiglas</td>
<td>1.51</td>
</tr>
<tr>
<td>Quartz, crystalline</td>
<td>1.544</td>
</tr>
<tr>
<td>Quartz, fused</td>
<td>1.458</td>
</tr>
<tr>
<td>Sodium chloride</td>
<td>1.544</td>
</tr>
<tr>
<td>Zircon</td>
<td>1.923</td>
</tr>
</tbody>
</table>

**Example 25.1 Speed of Light in Matter**

Calculate the speed of light in zircon, a material used in jewelry to imitate diamond.

**Strategy**

The speed of light in a material, \( v \), can be calculated from the index of refraction \( n \) of the material using the equation \( n = \frac{c}{v} \).

**Solution**

The equation for index of refraction states that \( n = \frac{c}{v} \). Rearranging this to determine \( v \) gives

\[
\begin{align*}
\frac{c}{n} &= v \\
25.5 & \\
\end{align*}
\]

The index of refraction for zircon is given as 1.923 in Table 25.1, and \( c \) is given in the equation for speed of light. Entering these values in the last expression gives

\[
\begin{align*}
v &= \frac{3.00 \times 10^{8} \text{ m/s}}{1.923} \\
&= 1.56 \times 10^{8} \text{ m/s}. \\
25.6 & \\
\end{align*}
\]

**Discussion**

This speed is slightly larger than half the speed of light in a vacuum and is still high compared with speeds we normally experience. The only substance listed in Table 25.1 that has a greater index of refraction than zircon is diamond. We shall see later that the large index of refraction for zircon makes it sparkle more than glass, but less than diamond.

**Law of Refraction**

Figure 25.12 shows how a ray of light changes direction when it passes from one medium to another. As before, the angles are measured relative to a perpendicular to the surface at the point where the light ray crosses it. (Some of the incident light will be reflected from the surface, but for now we will concentrate on the light that is transmitted.) The change in direction of the light ray depends on how the speed of light changes. The change in
the speed of light is related to the indices of refraction of the media involved. In the situations shown in Figure 25.12, medium 2 has a greater index of refraction than medium 1. This means that the speed of light is less in medium 2 than in medium 1. Note that as shown in Figure 25.12(a), the direction of the ray moves closer to the perpendicular when it slows down. Conversely, as shown in Figure 25.12(b), the direction of the ray moves away from the perpendicular when it speeds up. The path is exactly reversible. In both cases, you can imagine what happens by thinking about pushing a lawn mower from a footpath onto grass, and vice versa. Going from the footpath to grass, the front wheels are slowed and pulled to the side as shown. This is the same change in direction as for light when it goes from a fast medium to a slow one. When going from the grass to the footpath, the front wheels can move faster and the mower changes direction as shown. This, too, is the same change in direction as for light going from slow to fast.

Figure 25.12 The change in direction of a light ray depends on how the speed of light changes when it crosses from one medium to another. The speed of light is greater in medium 1 than in medium 2 in the situations shown here. (a) A ray of light moves closer to the perpendicular when it slows down. This is analogous to what happens when a lawn mower goes from a footpath to grass. (b) A ray of light moves away from the perpendicular when it speeds up. This is analogous to what happens when a lawn mower goes from grass to footpath. The paths are exactly reversible.

The amount that a light ray changes its direction depends both on the incident angle and the amount that the speed changes. For a ray at a given incident angle, a large change in speed causes a large change in direction, and thus a large change in angle. The exact mathematical relationship is the law of refraction, or “Snell’s Law,” which is stated in equation form as

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$  \hspace{1cm} (25.7)

Here $n_1$ and $n_2$ are the indices of refraction for medium 1 and 2, and $\theta_1$ and $\theta_2$ are the angles between the rays and the perpendicular in medium 1 and 2, as shown in Figure 25.12. The incoming ray is called the incident ray and the outgoing ray the refracted ray, and the associated angles the incident angle and the refracted angle. The law of refraction is also called Snell’s law after the Dutch mathematician Willebrord Snell (1591–1626), who discovered it in 1621. Snell’s experiments showed that the law of refraction was obeyed and that a characteristic index of refraction $n$ could be assigned to a given medium. Snell was not aware that the speed of light varied in different media, but through experiments he was able to determine indices of refraction from the way light rays changed direction.

The Law of Refraction

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$  \hspace{1cm} (25.8)

Take-Home Experiment: A Broken Pencil

A classic observation of refraction occurs when a pencil is placed in a glass half filled with water. Do this and observe the shape of the pencil when you look at the pencil sideways, that is, through air, glass, water. Explain your observations. Draw ray diagrams for the situation.

**Example 25.2 Determine the Index of Refraction from Refraction Data**

Find the index of refraction for medium 2 in Figure 25.12(a), assuming medium 1 is air and given the incident angle is 30.0º and the angle of refraction is 22.0º.

**Strategy**

The index of refraction for air is taken to be 1 in most cases (and up to four significant figures, it is 1.000). Thus $n_1 = 1.00$ here. From the given information, $\theta_1 = 30.0^\circ$ and $\theta_2 = 22.0^\circ$. With this information, the only unknown in Snell’s law is $n_2$, so that it can be used to find this unknown.

**Solution**

Snell’s law is

$$n_1 \sin \theta_1 = n_2 \sin \theta_2.$$  \hspace{1cm} (25.9)

Rearranging to isolate $n_2$ gives
\[ n_2 = n_1 \frac{\sin \theta_1}{\sin \theta_2} \]  

(25.10)

Entering known values,

\[ n_2 = \frac{1.00 \sin 30.0^\circ}{\sin 22.0^\circ} = \frac{0.500}{0.375} = 1.33. \]  

(25.11)

Discussion

This is the index of refraction for water, and Snell could have determined it by measuring the angles and performing this calculation. He would then have found 1.33 to be the appropriate index of refraction for water in all other situations, such as when a ray passes from water to glass. Today we can verify that the index of refraction is related to the speed of light in a medium by measuring that speed directly.

**Example 25.3 A Larger Change in Direction**

Suppose that in a situation like that in Example 25.2, light goes from air to diamond and that the incident angle is 30.0°. Calculate the angle of refraction \( \theta_2 \) in the diamond.

Strategy

Again the index of refraction for air is taken to be \( n_1 = 1.00 \), and we are given \( \theta_1 = 30.0^\circ \). We can look up the index of refraction for diamond in Table 25.1, finding \( n_2 = 2.419 \). The only unknown in Snell’s law is \( \theta_2 \), which we wish to determine.

Solution

Solving Snell’s law for \( \sin \theta_2 \) yields

\[ \sin \theta_2 = \frac{n_1}{n_2} \sin \theta_1. \]  

(25.12)

Entering known values,

\[ \sin \theta_2 = \frac{1.00 \sin 30.0^\circ}{2.419} = 0.500 \times (0.413) = 0.207. \]  

(25.13)

The angle is thus

\[ \theta_2 = \sin^{-1} 0.207 = 11.9^\circ. \]  

(25.14)

Discussion

For the same 30° angle of incidence, the angle of refraction in diamond is significantly smaller than in water (11.9° rather than 22°—see the preceding example). This means there is a larger change in direction in diamond. The cause of a large change in direction is a large change in the index of refraction (or speed). In general, the larger the change in speed, the greater the effect on the direction of the ray.

### 25.4 Total Internal Reflection

A good-quality mirror may reflect more than 90% of the light that falls on it, absorbing the rest. But it would be useful to have a mirror that reflects all of the light that falls on it. Interestingly, we can produce total reflection using an aspect of refraction.

Consider what happens when a ray of light strikes the surface between two materials, such as is shown in Figure 25.13(a). Part of the light crosses the boundary and is refracted; the rest is reflected. If, as shown in the figure, the index of refraction for the second medium is less than for the first, the ray bends away from the perpendicular. (Since \( n_1 > n_2 \), the angle of refraction is greater than the angle of incidence—that is, \( \theta_2 \) is greater than \( \theta_1 \).) Now imagine what happens as the incident angle is increased. This causes \( \theta_2 \) to increase also. The largest angle of refraction \( \theta_2 \) can be is 90°, as shown in Figure 25.13(b). The critical angle \( \theta_c \) for a combination of materials is defined to be the incident angle \( \theta_1 \) that produces an angle of refraction of 90°. That is, \( \theta_c \) is the incident angle for which \( \theta_2 = 90^\circ \). If the incident angle \( \theta_1 \) is greater than the critical angle, as shown in Figure 25.13(c), then all of the light is reflected back into medium 1, a condition called total internal reflection.

**Critical Angle**

The incident angle \( \theta_1 \) that produces an angle of refraction of 90° is called the critical angle, \( \theta_c \).
Figure 25.13 (a) A ray of light crosses a boundary where the speed of light increases and the index of refraction decreases. That is, \( n_2 < n_1 \). The ray bends away from the perpendicular. (b) The critical angle \( \theta_c \) is the one for which the angle of refraction is \( 90^\circ \). (c) Total internal reflection occurs when the incident angle is greater than the critical angle.

Snell's law states the relationship between angles and indices of refraction. It is given by

\[
 n_1 \sin \theta_1 = n_2 \sin \theta_2. \tag{25.15}
\]

When the incident angle equals the critical angle \( \theta_1 = \theta_c \), the angle of refraction is \( 90^\circ \) \( \theta_2 = 90^\circ \). Noting that \( \sin 90^\circ = 1 \), Snell's law in this case becomes

\[
 n_1 \sin \theta_1 = n_2. \tag{25.16}
\]

The critical angle \( \theta_c \) for a given combination of materials is thus

\[
 \theta_c = \sin^{-1}(n_2 / n_1) \text{ for } n_1 > n_2. \tag{25.17}
\]

Total internal reflection occurs for any incident angle greater than the critical angle \( \theta_c \), and it can only occur when the second medium has an index of refraction less than the first. Note the above equation is written for a light ray that travels in medium 1 and reflects from medium 2, as shown in the figure.

**Example 25.4 How Big is the Critical Angle Here?**

What is the critical angle for light traveling in a polystyrene (a type of plastic) pipe surrounded by air?

**Strategy**
The index of refraction for polystyrene is found to be 1.49 in Figure 25.14, and the index of refraction of air can be taken to be 1.00, as before. Thus, the condition that the second medium (air) has an index of refraction less than the first (plastic) is satisfied, and the equation

\[ \theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \]

can be used to find the critical angle \( \theta_c \). Here, then, \( n_2 = 1.00 \) and \( n_1 = 1.49 \).

**Solution**

The critical angle is given by

\[ \theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right). \]  \hspace{1cm} (25.18)

Substituting the identified values gives

\[ \theta_c = \sin^{-1}\left(\frac{1.00}{1.49}\right) = \sin^{-1}(0.671) = 42.2^\circ. \]  \hspace{1cm} (25.19)

**Discussion**

This means that any ray of light inside the plastic that strikes the surface at an angle greater than \( 42.2^\circ \) will be totally reflected. This will make the inside surface of the clear plastic a perfect mirror for such rays without any need for the silvering used on common mirrors. Different combinations of materials have different critical angles, but any combination with \( n_1 > n_2 \) can produce total internal reflection. The same calculation as made here shows that the critical angle for a ray going from water to air is \( 48.6^\circ \), while that from diamond to air is \( 24.4^\circ \), and that from flint glass to crown glass is \( 66.3^\circ \). There is no total reflection for rays going in the other direction—for example, from air to water—since the condition that the second medium must have a smaller index of refraction is not satisfied. A number of interesting applications of total internal reflection follow.

**Fiber Optics: Endoscopes to Telephones**

Fiber optics is one application of total internal reflection that is in wide use. In communications, it is used to transmit telephone, internet, and cable TV signals. **Fiber optics** employs the transmission of light down fibers of plastic or glass. Because the fibers are thin, light entering one is likely to strike the inside surface at an angle greater than the critical angle and, thus, be totally reflected (See Figure 25.14.) The index of refraction outside the fiber must be smaller than inside, a condition that is easily satisfied by coating the outside of the fiber with a material having an appropriate refractive index. In fact, most fibers have a varying refractive index to allow more light to be guided along the fiber through total internal refraction. Rays are reflected around corners as shown, making the fibers into tiny light pipes.

**Figure 25.14** Light entering a thin fiber may strike the inside surface at large or grazing angles and is completely reflected if these angles exceed the critical angle. Such rays continue down the fiber, even following it around corners, since the angles of reflection and incidence remain large.

Bundles of fibers can be used to transmit an image without a lens, as illustrated in Figure 25.15. The output of a device called an **endoscope** is shown in Figure 25.15(b). Endoscopes are used to explore the body through various orifices or minor incisions. Light is transmitted down one fiber bundle to illuminate internal parts, and the reflected light is transmitted back out through another to be observed. Surgery can be performed, such as arthroscopic surgery on the knee joint, employing cutting tools attached to and observed with the endoscope. Samples can also be obtained, such as by lassoing an intestinal polyp for external examination.

Fiber optics has revolutionized surgical techniques and observations within the body. There are a host of medical diagnostic and therapeutic uses. The flexibility of the fiber optic bundle allows it to navigate around difficult and small regions in the body, such as the intestines, the heart, blood vessels, and joints. Transmission of an intense laser beam to burn away obstructing plaques in major arteries as well as delivering light to activate chemotherapy drugs are becoming commonplace. Optical fibers have in fact enabled microsurgery and remote surgery where the incisions are small and the surgeon’s fingers do not need to touch the diseased tissue.
Fibers in bundles are surrounded by a cladding material that has a lower index of refraction than the core. (See Figure 25.16.) The cladding prevents light from being transmitted between fibers in a bundle. Without cladding, light could pass between fibers in contact, since their indices of refraction are identical. Since no light gets into the cladding (there is total internal reflection back into the core), none can be transmitted between clad fibers that are in contact with one another. The cladding prevents light from escaping out of the fiber; instead most of the light is propagated along the length of the fiber, minimizing the loss of signal and ensuring that a quality image is formed at the other end. The cladding and an additional protective layer make optical fibers flexible and durable.

Special tiny lenses that can be attached to the ends of bundles of fibers are being designed and fabricated. Light emerging from a fiber bundle can be focused and a tiny spot can be imaged. In some cases the spot can be scanned, allowing quality imaging of a region inside the body. Special minute optical filters inserted at the end of the fiber bundle have the capacity to image tens of microns below the surface without cutting the surface—non-intrusive diagnostics. This is particularly useful for determining the extent of cancers in the stomach and bowel.

Most telephone conversations and Internet communications are now carried by laser signals along optical fibers. Extensive optical fiber cables have been placed on the ocean floor and underground to enable optical communications. Optical fiber communication systems offer several advantages over electrical (copper) based systems, particularly for long distances. The fibers can be made so transparent that light can travel many kilometers before it becomes dim enough to require amplification—much superior to copper conductors. This property of optical fibers is called low loss. Lasers emit light with characteristics that allow far more conversations in one fiber than are possible with electric signals on a single conductor. This property of optical fibers is called high bandwidth. Optical signals in one fiber do not produce undesirable effects in other adjacent fibers. This property of optical fibers is called reduced crosstalk. We shall explore the unique characteristics of laser radiation in a later chapter.

Corner Reflectors and Diamonds

A light ray that strikes an object consisting of two mutually perpendicular reflecting surfaces is reflected back exactly parallel to the direction from which it came. This is true whenever the reflecting surfaces are perpendicular, and it is independent of the angle of incidence. Such an object, shown in Figure 25.52, is called a corner reflector, since the light bounces from its inside corner. Many inexpensive reflector buttons on bicycles, cars, and warning signs have corner reflectors designed to return light in the direction from which it originated. It was more expensive for astronauts to place one on the moon. Laser signals can be bounced from that corner reflector to measure the gradually increasing distance to the moon with great precision.
Corner reflectors are perfectly efficient when the conditions for total internal reflection are satisfied. With common materials, it is easy to obtain a critical angle that is less than 45°. One use of these perfect mirrors is in binoculars, as shown in Figure 25.18. Another use is in periscopes found in submarines.

The Sparkle of Diamonds

Total internal reflection, coupled with a large index of refraction, explains why diamonds sparkle more than other materials. The critical angle for a diamond-to-air surface is only 24.4°, and so when light enters a diamond, it has trouble getting back out. (See Figure 25.19.) Although light freely enters the diamond, it can exit only if it makes an angle less than 24.4°. Facets on diamonds are specifically intended to make this unlikely, so that the light can exit only in certain places. Good diamonds are very clear, so that the light makes many internal reflections and is concentrated at the few places it can exit—hence the sparkle. (Zircon is a natural gemstone that has an exceptionally large index of refraction, but not as large as diamond, so it is not as highly prized. Cubic zirconia is manufactured and has an even higher index of refraction (≈ 2.17), but still less than that of diamond.) The colors you see emerging from a sparkling diamond are not due to the diamond’s color, which is usually nearly colorless. Those colors result from dispersion, the topic of Dispersion: The Rainbow and Prisms. Colored diamonds get their color from structural defects of the crystal lattice and the inclusion of minute quantities of graphite and other materials. The Argyle Mine in Western Australia produces around 90% of the world’s pink, red, champagne, and cognac diamonds, while around 50% of the world’s clear diamonds come from central and southern Africa.
Figure 25.19 Light cannot easily escape a diamond, because its critical angle with air is so small. Most reflections are total, and the facets are placed so that light can exit only in particular ways—thus concentrating the light and making the diamond sparkle.

PhET Explorations: Bending Light
Explore bending of light between two media with different indices of refraction. See how changing from air to water to glass changes the bending angle. Play with prisms of different shapes and make rainbows.

Figure 25.20 Bending Light (http://cnx.org/content/m42462/1.5/bending-light_en.jar)

25.5 Dispersion: The Rainbow and Prisms
Everyone enjoys the spectacle of a rainbow glimmering against a dark stormy sky. How does sunlight falling on clear drops of rain get broken into the rainbow of colors we see? The same process causes white light to be broken into colors by a clear glass prism or a diamond. (See Figure 25.21.)

Figure 25.21 The colors of the rainbow (a) and those produced by a prism (b) are identical. (credit: Alfredo55, Wikimedia Commons; NASA)

We see about six colors in a rainbow—red, orange, yellow, green, blue, and violet; sometimes indigo is listed, too. Those colors are associated with different wavelengths of light, as shown in Figure 25.22. When our eye receives pure-wavelength light, we tend to see only one of the six colors, depending on wavelength. The thousands of other hues we can sense in other situations are our eye’s response to various mixtures of wavelengths. White light, in particular, is a fairly uniform mixture of all visible wavelengths. Sunlight, considered to be white, actually appears to be a bit yellow because of its mixture of wavelengths, but it does contain all visible wavelengths. The sequence of colors in rainbows is the same sequence as the colors plotted versus wavelength in Figure 25.22. What this implies is that white light is spread out according to wavelength in a rainbow. Dispersion is defined as the spreading of white light into its full spectrum of wavelengths. More technically, dispersion occurs whenever there is a process that changes the direction of light in a manner that depends on wavelength. Dispersion, as a general phenomenon, can occur for any type of wave and always involves wavelength-dependent processes.
Dispersion

Dispersion is defined to be the spreading of white light into its full spectrum of wavelengths.

Figure 25.22 Even though rainbows are associated with seven colors, the rainbow is a continuous distribution of colors according to wavelengths.

Refraction is responsible for dispersion in rainbows and many other situations. The angle of refraction depends on the index of refraction, as we saw in The Law of Refraction. We know that the index of refraction $n$ depends on the medium. But for a given medium, $n$ also depends on wavelength. (See Table 25.2. Note that, for a given medium, $n$ increases as wavelength decreases and is greatest for violet light. Thus violet light is bent more than red light, as shown for a prism in Figure 25.23(b), and the light is dispersed into the same sequence of wavelengths as seen in Figure 25.21 and Figure 25.22.

Making Connections: Dispersion

Any type of wave can exhibit dispersion. Sound waves, all types of electromagnetic waves, and water waves can be dispersed according to wavelength. Dispersion occurs whenever the speed of propagation depends on wavelength, thus separating and spreading out various wavelengths. Dispersion may require special circumstances and can result in spectacular displays such as in the production of a rainbow. This is also true for sound, since all frequencies ordinarily travel at the same speed. If you listen to sound through a long tube, such as a vacuum cleaner hose, you can easily hear it is dispersed by interaction with the tube. Dispersion, in fact, can reveal a great deal about what the wave has encountered that disperses its wavelengths. The dispersion of electromagnetic radiation from outer space, for example, has revealed much about what exists between the stars—the so-called empty space.

Table 25.2 Index of Refraction $n$ in Selected Media at Various Wavelengths

<table>
<thead>
<tr>
<th>Medium</th>
<th>Red (660 nm)</th>
<th>Orange (610 nm)</th>
<th>Yellow (580 nm)</th>
<th>Green (550 nm)</th>
<th>Blue (470 nm)</th>
<th>Violet (410 nm)</th>
</tr>
</thead>
</table>
Rainbows are produced by a combination of refraction and reflection. You may have noticed that you see a rainbow only when you look away from the sun. Light enters a drop of water and is reflected from the back of the drop, as shown in Figure 25.24. The light is refracted both as it enters and as it leaves the drop. Since the index of refraction of water varies with wavelength, the light is dispersed, and a rainbow is observed, as shown in Figure 25.25 (a). (There is no dispersion caused by reflection at the back surface, since the law of reflection does not depend on wavelength.) The actual rainbow of colors seen by an observer depends on the myriad of rays being refracted and reflected toward the observer’s eyes from numerous drops of water. The effect is most spectacular when the background is dark, as in stormy weather, but can also be observed in waterfalls and lawn sprinklers. The arc of a rainbow comes from the need to be looking at a specific angle relative to the direction of the sun, as illustrated in Figure 25.25 (b). (If there are two reflections of light within the water drop, another “secondary” rainbow is produced. This rare event produces an arc that lies above the primary rainbow arc—see Figure 25.25 (c).)
Figure 25.25 (a) Different colors emerge in different directions, and so you must look at different locations to see the various colors of a rainbow. (b) The arc of a rainbow results from the fact that a line between the observer and any point on the arc must make the correct angle with the parallel rays of sunlight to receive the refracted rays. (c) Double rainbow. (credit: Nicholas, Wikimedia Commons)

Dispersion may produce beautiful rainbows, but it can cause problems in optical systems. White light used to transmit messages in a fiber is dispersed, spreading out in time and eventually overlapping with other messages. Since a laser produces a nearly pure wavelength, its light experiences little dispersion, an advantage over white light for transmission of information. In contrast, dispersion of electromagnetic waves coming to us from outer space can be used to determine the amount of matter they pass through. As with many phenomena, dispersion can be useful or a nuisance, depending on the situation and our human goals.

PhET Explorations: Geometric Optics

How does a lens form an image? See how light rays are refracted by a lens. Watch how the image changes when you adjust the focal length of the lens, move the object, move the lens, or move the screen.
25.6 Image Formation by Lenses

Lenses are found in a huge array of optical instruments, ranging from a simple magnifying glass to the eye to a camera’s zoom lens. In this section, we will use the law of refraction to explore the properties of lenses and how they form images.

The word lens derives from the Latin word for a lentil bean, the shape of which is similar to the convex lens in Figure 25.27. The convex lens shown has been shaped so that all light rays that enter it parallel to its axis cross one another at a single point on the opposite side of the lens. (The axis is defined to be a line normal to the lens at its center, as shown in Figure 25.27.) Such a lens is called a converging (or convex) lens for the converging effect it has on light rays. An expanded view of the path of one ray through the lens is shown, to illustrate how the ray changes direction both as it enters and as it leaves the lens. Since the index of refraction of the lens is greater than that of air, the ray moves towards the perpendicular as it enters and away from the perpendicular as it leaves. (This is in accordance with the law of refraction.) Due to the lens’s shape, light is thus bent toward the axis at both surfaces. The point at which the rays cross is defined to be the focal point \( F \) of the lens. The distance from the center of the lens to its focal point is defined to be the focal length \( f \) of the lens. Figure 25.28 shows how a converging lens, such as that in a magnifying glass, can converge the nearly parallel light rays from the sun to a small spot.

**Converging or Convex Lens**

The lens in which light rays that enter it parallel to its axis cross one another at a single point on the opposite side with a converging effect is called converging lens.

**Focal Point \( F \)**

The point at which the light rays cross is called the focal point \( F \) of the lens.

**Focal Length \( f \)**

The distance from the center of the lens to its focal point is called focal length \( f \).
Figure 25.28 Sunlight focused by a converging magnifying glass can burn paper. Light rays from the sun are nearly parallel and cross at the focal point of the lens. The more powerful the lens, the closer to the lens the rays will cross.

The greater effect a lens has on light rays, the more powerful it is said to be. For example, a powerful converging lens will focus parallel light rays closer to itself and will have a smaller focal length than a weak lens. The light will also focus into a smaller and more intense spot for a more powerful lens. The power $P$ of a lens is defined to be the inverse of its focal length. In equation form, this is

$$ P = \frac{1}{f} \quad \text{(25.20)} $$

where $f$ is the focal length of the lens, which must be given in meters (and not cm or mm). The power of a lens $P$ has the unit diopters (D), provided that the focal length is given in meters. That is, 1 D = $1/m$, or $1 \text{ m}^{-1}$. (Note that this power (optical power, actually) is not the same as power in watts defined in Work, Energy, and Energy Resources. It is a concept related to the effect of optical devices on light.) Optometrists prescribe common spectacles and contact lenses in units of diopters.

Example 25.5 What is the Power of a Common Magnifying Glass?

Suppose you take a magnifying glass out on a sunny day and you find that it concentrates sunlight to a small spot 8.00 cm away from the lens. What are the focal length and power of the lens?

**Strategy**

The situation here is the same as those shown in Figure 25.27 and Figure 25.28. The Sun is so far away that the Sun's rays are nearly parallel when they reach Earth. The magnifying glass is a convex (or converging) lens, focusing the nearly parallel rays of sunlight. Thus the focal length of the lens is the distance from the lens to the spot, and its power is the inverse of this distance (in m).

**Solution**

The focal length of the lens is the distance from the center of the lens to the spot, given to be 8.00 cm. Thus,

$$ f = 8.00 \text{ cm}. \quad \text{(25.22)} $$

To find the power of the lens, we must first convert the focal length to meters; then, we substitute this value into the equation for power. This gives

$$ P = \frac{1}{f} = \frac{1}{0.0800 \text{ m}} = 12.5 \text{ D}. \quad \text{(25.23)} $$

**Discussion**

This is a relatively powerful lens. The power of a lens in diopters should not be confused with the familiar concept of power in watts. It is an unfortunate fact that the word “power” is used for two completely different concepts. If you examine a prescription for eyeglasses, you will note lens powers given in diopters. If you examine the label on a motor, you will note energy consumption rate given as power in watts.

Figure 25.29 shows a concave lens and the effect it has on rays of light that enter it parallel to its axis (the path taken by ray 2 in the figure is the axis of the lens). The concave lens is a diverging lens, because it causes the light rays to bend away (diverge) from its axis. In this case, the lens has been shaped so that all light rays entering it parallel to its axis appear to originate from the same point, $F'$, defined to be the focal point of a diverging lens. The distance from the center of the lens to the focal point is again called the focal length $f'$ of the lens. Note that the focal length and power of
A diverging lens are defined to be negative. For example, if the distance to \( F \) in Figure 25.29 is 5.00 cm, then the focal length is \( f = -5.00 \text{ cm} \) and the power of the lens is \( P = -20 \text{ D} \). An expanded view of the path of one ray through the lens is shown in the figure to illustrate how the shape of the lens, together with the law of refraction, causes the ray to follow its particular path and be diverged.

**Figure 25.29** Rays of light entering a diverging lens parallel to its axis are diverged, and all appear to originate at its focal point \( F \). The dashed lines are not rays—they indicate the directions from which the rays appear to come. The focal length \( f \) of a diverging lens is negative. An expanded view of the path taken by ray 1 shows the perpendiculars and the angles of incidence and refraction at both surfaces.

**Diverging Lens**

A lens that causes the light rays to bend away from its axis is called a diverging lens.

As noted in the initial discussion of the law of refraction in The Law of Refraction, the paths of light rays are exactly reversible. This means that the direction of the arrows could be reversed for all of the rays in Figure 25.27 and Figure 25.29. For example, if a point light source is placed at the focal point of a convex lens, as shown in Figure 25.30, parallel light rays emerge from the other side.

**Figure 25.30** A small light source, like a light bulb filament, placed at the focal point of a convex lens, results in parallel rays of light emerging from the other side. The paths are exactly the reverse of those shown in Figure 25.27. This technique is used in lighthouses and sometimes in traffic lights to produce a directional beam of light from a source that emits light in all directions.

**Ray Tracing and Thin Lenses**

Ray tracing is the technique of determining or following (tracing) the paths that light rays take. For rays passing through matter, the law of refraction is used to trace the paths. Here we use ray tracing to help us understand the action of lenses in situations ranging from forming images on film to magnifying small print to correcting nearsightedness. While ray tracing for complicated lenses, such as those found in sophisticated cameras, may require computer techniques, there is a set of simple rules for tracing rays through thin lenses. A thin lens is defined to be one whose thickness allows rays to refract, as illustrated in Figure 25.27, but does not allow properties such as dispersion and aberrations. An ideal thin lens has two refracting surfaces but the lens is thin enough to assume that light rays bend only once. A thin symmetrical lens has two focal points, one on either side and both at the same distance from the lens. (See Figure 25.31.) Another important characteristic of a thin lens is that light rays through its center are deflected by a negligible amount, as seen in Figure 25.32.

**Thin Lens**

A thin lens is defined to be one whose thickness allows rays to refract but does not allow properties such as dispersion and aberrations.

**Take-Home Experiment: A Visit to the Optician**

Look through your eyeglasses (or those of a friend) backward and forward and comment on whether they act like thin lenses.
Figure 25.31 Thin lenses have the same focal length on either side. (a) Parallel light rays entering a converging lens from the right cross at its focal point on the left. (b) Parallel light rays entering a diverging lens from the right seem to come from the focal point on the right.

Figure 25.32 The light ray through the center of a thin lens is deflected by a negligible amount and is assumed to emerge parallel to its original path (shown as a shaded line).

Using paper, pencil, and a straight edge, ray tracing can accurately describe the operation of a lens. The rules for ray tracing for thin lenses are based on the illustrations already discussed:

1. A ray entering a converging lens parallel to its axis passes through the focal point $F$ of the lens on the other side. (See rays 1 and 3 in Figure 25.27.)
2. A ray entering a diverging lens parallel to its axis seems to come from the focal point $F$. (See rays 1 and 3 in Figure 25.29.)
3. A ray passing through the center of either a converging or a diverging lens does not change direction. (See Figure 25.32, and see ray 2 in Figure 25.27 and Figure 25.29.)
4. A ray entering a converging lens through its focal point exits parallel to its axis. (The reverse of rays 1 and 3 in Figure 25.27.)
5. A ray that enters a diverging lens by heading toward the focal point on the opposite side exits parallel to the axis. (The reverse of rays 1 and 3 in Figure 25.29.)

**Rules for Ray Tracing**

1. A ray entering a converging lens parallel to its axis passes through the focal point $F$ of the lens on the other side.
2. A ray entering a diverging lens parallel to its axis seems to come from the focal point $F$.
3. A ray passing through the center of either a converging or a diverging lens does not change direction.
4. A ray entering a converging lens through its focal point exits parallel to its axis.
5. A ray that enters a diverging lens by heading toward the focal point on the opposite side exits parallel to the axis.
Image Formation by Thin Lenses

In some circumstances, a lens forms an obvious image, such as when a movie projector casts an image onto a screen. In other cases, the image is less obvious. Where, for example, is the image formed by eyeglasses? We will use ray tracing for thin lenses to illustrate how they form images, and we will develop equations to describe the image formation quantitatively.

Consider an object some distance away from a converging lens, as shown in Figure 25.33. To find the location and size of the image formed, we trace the paths of selected light rays originating from one point on the object, in this case the top of the person’s head. The figure shows three rays from the top of the object that can be traced using the ray tracing rules given above. (Rays leave this point going in many directions, but we concentrate on only a few with paths that are easy to trace.) The first ray is one that enters the lens parallel to its axis and passes through the focal point on the other side (rule 1). The second ray passes through the center of the lens without changing direction (rule 3). The third ray passes through the nearer focal point on its way into the lens and leaves the lens parallel to its axis (rule 4). The three rays cross at the same point on the other side of the lens. The image of the top of the person’s head is located at this point. All rays that come from the same point on the top of the person’s head are refracted in such a way as to cross at the point shown. Rays from another point on the object, such as her belt buckle, will also cross at another common point, forming a complete image, as shown. Although three rays are traced in Figure 25.33, only two are necessary to locate the image. It is best to trace rays for which there are simple ray tracing rules. Before applying ray tracing to other situations, let us consider the example shown in Figure 25.33 in more detail.

**Figure 25.33** Ray tracing is used to locate the image formed by a lens. Rays originating from the same point on the object are traced—the three chosen rays each follow one of the rules for ray tracing, so that their paths are easy to determine. The image is located at the point where the rays cross. In this case, a real image—one that can be projected on a screen—is formed.

The image formed in Figure 25.33 is a **real image**, meaning that it can be projected. That is, light rays from one point on the object actually cross at the location of the image and can be projected onto a screen, a piece of film, or the retina of an eye, for example. Figure 25.34 shows how such an image would be projected onto film by a camera lens. This figure also shows how a real image is projected onto the retina by the lens of an eye. Note that the image is there whether it is projected onto a screen or not.

**Real Image**

The image in which light rays from one point on the object actually cross at the location of the image and can be projected onto a screen, a piece of film, or the retina of an eye is called a real image.
Several important distances appear in Figure 25.33. We define $d_o$ to be the object distance, the distance of an object from the center of a lens. Image distance $d_i$ is defined to be the distance of the image from the center of a lens. The height of the object and height of the image are given the symbols $h_o$ and $h_i$, respectively. Images that appear upright relative to the object have heights that are positive and those that are inverted have negative heights. Using the rules of ray tracing and making a scale drawing with paper and pencil, like that in Figure 25.33, we can accurately describe the location and size of an image. But the real benefit of ray tracing is in visualizing how images are formed in a variety of situations. To obtain numerical information, we use a pair of equations that can be derived from a geometric analysis of ray tracing for thin lenses. The thin lens equations are

\begin{align}
\frac{1}{d_o} + \frac{1}{d_i} &= \frac{1}{f} \\
\frac{h_i}{h_o} &= -\frac{d_i}{d_o} = m.
\end{align}

We define the ratio of image height to object height ($h_i / h_o$) to be the magnification $m$. (The minus sign in the equation above will be discussed shortly.) The thin lens equations are broadly applicable to all situations involving thin lenses (and “thin” mirrors, as we will see later). We will explore many features of image formation in the following worked examples.

### Example 25.6 Finding the Image of a Light Bulb Filament by Ray Tracing and by the Thin Lens Equations

A clear glass light bulb is placed 0.750 m from a convex lens having a 0.500 m focal length, as shown in Figure 25.35. Use ray tracing to get an approximate location for the image. Then use the thin lens equations to calculate (a) the location of the image and (b) its magnification. Verify that ray tracing and the thin lens equations produce consistent results.
A light bulb placed 0.750 m from a lens having a 0.500 m focal length produces a real image on a poster board as discussed in the example above. Ray tracing predicts the image location and size.

**Strategy and Concept**

Since the object is placed farther away from a converging lens than the focal length of the lens, this situation is analogous to those illustrated in Figure 25.33 and Figure 25.34. Ray tracing to scale should produce similar results for $d_i$. Numerical solutions for $d_i$ and $m$ can be obtained using the thin lens equations, noting that $d_o = 0.750$ m and $f = 0.500$ m.

**Solutions (Ray tracing)**

The ray tracing to scale in Figure 25.35 shows two rays from a point on the bulb’s filament crossing about 1.50 m on the far side of the lens. Thus the image distance $d_i$ is about 1.50 m. Similarly, the image height based on ray tracing is greater than the object height by about a factor of 2, and the image is inverted. Thus $m$ is about –2. The minus sign indicates that the image is inverted.

The thin lens equations can be used to find $d_i$ from the given information:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}. \tag{25.28}$$

Rearranging to isolate $d_i$ gives

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}. \tag{25.29}$$

Entering known quantities gives a value for $1/d_i$:

$$\frac{1}{d_i} = \frac{1}{0.500 \text{ m}} - \frac{1}{0.750 \text{ m}} = 0.667 \text{ m}. \tag{25.30}$$

This must be inverted to find $d_i$:

$$d_i = \frac{m}{0.667} = 1.50 \text{ m}. \tag{25.31}$$

Note that another way to find $d_i$ is to rearrange the equation:

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}. \tag{25.32}$$

This yields the equation for the image distance as:

$$d_i = \frac{fd_o}{d_o - f}. \tag{25.33}$$

Note that there is no inverting here.

The thin lens equations can be used to find the magnification $m$, since both $d_i$ and $d_o$ are known. Entering their values gives

$$m = \frac{-d_i}{d_o} = -\frac{1.50 \text{ m}}{0.750 \text{ m}} = -2.00. \tag{25.34}$$

**Discussion**

Note that the minus sign causes the magnification to be negative when the image is inverted. Ray tracing and the use of the thin lens equations produce consistent results. The thin lens equations give the most precise results, being limited only by the accuracy of the given information. Ray tracing is limited by the accuracy with which you can draw, but it is highly useful both conceptually and visually.

Real images, such as the one considered in the previous example, are formed by converging lenses whenever an object is farther from the lens than its focal length. This is true for movie projectors, cameras, and the eye. We shall refer to these as case 1 images. A case 1 image is formed when $d_o > f$ and $f$ is positive, as in Figure 25.36(a). (A summary of the three cases or types of image formation appears at the end of this section.)
A different type of image is formed when an object, such as a person's face, is held close to a convex lens. The image is upright and larger than the object, as seen in Figure 25.36(b), and so the lens is called a magnifier. If you slowly pull the magnifier away from the face, you will see that the magnification steadily increases until the image begins to blur. Pulling the magnifier even farther away produces an inverted image as seen in Figure 25.36(a). The distance at which the image blurs, and beyond which it inverts, is the focal length of the lens. To use a convex lens as a magnifier, the object must be closer to the converging lens than its focal length. This is called a case 2 image. A case 2 image is formed when \( d_o < f \) and \( f \) is positive.

![Figure 25.36](image)

**Figure 25.36** (a) When a converging lens is held farther away from the face than the lens's focal length, an inverted image is formed. This is a case 1 image. Note that the image is in focus but the face is not, because the image is much closer to the camera taking this photograph than the face. (credit: DaMongMan, Flickr) (b) A magnified image of a face is produced by placing it closer to the converging lens than its focal length. This is a case 2 image. (credit: Casey Fleser, Flickr)

**Figure 25.37** uses ray tracing to show how an image is formed when an object is held closer to a converging lens than its focal length. Rays coming from a common point on the object continue to diverge after passing through the lens, but all appear to originate from a point at the location of the image. The image is on the same side of the lens as the object and is farther away from the lens than the object. This image, like all case 2 images, cannot be projected and, hence, is called a virtual image. Light rays only appear to originate at a virtual image: they do not actually pass through that location in space. A screen placed at the location of a virtual image will receive only diffuse light from the object, not focused rays from the lens. Additionally, a screen placed on the opposite side of the lens will receive rays that are still diverging, and so no image will be projected on it. We can see the magnified image with our eyes, because the lens of the eye converges the rays into a real image projected on our retina. Finally, we note that a virtual image is upright and larger than the object, meaning that the magnification is positive and greater than 1.
Virtual Image

An image that is on the same side of the lens as the object and cannot be projected on a screen is called a virtual image.

Example 25.7 Image Produced by a Magnifying Glass

Suppose the book page in Figure 25.37 (a) is held 7.50 cm from a convex lens of focal length 10.0 cm, such as a typical magnifying glass might have. What magnification is produced?

Strategy and Concept

We are given that \( d_o = 7.50 \text{ cm} \) and \( f = 10.0 \text{ cm} \), so we have a situation where the object is placed closer to the lens than its focal length. We therefore expect to get a case 2 virtual image with a positive magnification that is greater than 1. Ray tracing produces an image like that shown in Figure 25.37, but we will use the thin lens equations to get numerical solutions in this example.

Solution

To find the magnification \( m \), we try to use magnification equation, \( m = -d_i/d_o \). We do not have a value for \( d_i \), so that we must first find the location of the image using lens equation. (The procedure is the same as followed in the preceding example, where \( d_o \) and \( f \) were known.)

Rearranging the magnification equation to isolate \( d_i \) gives

\[
\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}.
\]  

(25.35)

Entering known values, we obtain a value for \( 1/d_i \):

\[
\frac{1}{d_i} = \frac{1}{10.0 \text{ cm}} - \frac{1}{7.50 \text{ cm}} = -0.0333 \text{ cm}^{-1}.
\]  

(25.36)

This must be inverted to find \( d_i \):

\[
d_i = -\frac{0.0333 \text{ cm}}{1} = -30.0 \text{ cm}.
\]  

(25.37)

Now the thin lens equation can be used to find the magnification \( m \), since both \( d_i \) and \( d_o \) are known. Entering their values gives

\[
m = -\frac{d_i}{d_o} = -\frac{-30.0 \text{ cm}}{10.0 \text{ cm}} = 3.00.
\]  

(25.38)
Discussion

A number of results in this example are true of all case 2 images, as well as being consistent with Figure 25.37. Magnification is indeed positive (as predicted), meaning the image is upright. The magnification is also greater than 1, meaning that the image is larger than the object—in this case, by a factor of 3. Note that the image distance is negative. This means the image is on the same side of the lens as the object. Thus the image cannot be projected and is virtual. (Negative values of $d_i$ occur for virtual images.) The image is farther from the lens than the object, since the image distance is greater in magnitude than the object distance. The location of the image is not obvious when you look through a magnifier. In fact, since the image is bigger than the object, you may think the image is closer than the object. But the image is farther away, a fact that is useful in correcting farsightedness, as we shall see in a later section.

A third type of image is formed by a diverging or concave lens. Try looking through eyeglasses meant to correct nearsightedness. (See Figure 25.38.) You will see an image that is upright but smaller than the object. This means that the magnification is positive but less than 1. The ray diagram in Figure 25.39 shows that the image is on the same side of the lens as the object and, hence, cannot be projected—it is a virtual image. Note that the image is closer to the lens than the object. This is a case 3 image, formed for any object by a negative focal length or diverging lens.

Example 25.8 Image Produced by a Concave Lens

Suppose an object such as a book page is held 7.50 cm from a concave lens of focal length $-10.0$ cm. Such a lens could be used in eyeglasses to correct pronounced nearsightedness. What magnification is produced?

Strategy and Concept

This example is identical to the preceding one, except that the focal length is negative for a concave or diverging lens. The method of solution is thus the same, but the results are different in important ways.
Solution
To find the magnification \( m \), we must first find the image distance \( d_i \) using thin lens equation

\[
\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o},
\]

or its alternative rearrangement

\[
d_i = \frac{f d_o}{d_o - f}.
\]

We are given that \( f = -10.0 \text{ cm} \) and \( d_o = 7.50 \text{ cm} \). Entering these yields a value for \( 1/d_i \):

\[
\frac{1}{d_i} = -\frac{1}{10.0 \text{ cm}} - \frac{1}{7.50 \text{ cm}} = -\frac{0.2333}{\text{cm}}.
\]

This must be inverted to find \( d_i \):

\[
d_i = -\frac{\text{cm}}{0.2333} = -4.29 \text{ cm}.
\]

Or

\[
d_i = \frac{(7.5)(-10)}{(7.5 - (-10))} = -75 / 17.5 = -4.29 \text{ cm}.
\]

Now the magnification equation can be used to find the magnification \( m \), since both \( d_i \) and \( d_o \) are known. Entering their values gives

\[
m = -\frac{d_i}{d_o} = -\frac{-4.29 \text{ cm}}{7.50 \text{ cm}} = 0.571.
\]

Discussion
A number of results in this example are true of all case 3 images, as well as being consistent with Figure 25.39. Magnification is positive (as predicted), meaning the image is upright. The magnification is also less than 1, meaning the image is smaller than the object—in this case, a little over half its size. The image distance is negative, meaning the image is on the same side of the lens as the object. (The image is virtual.) The image is closer to the lens than the object, since the image distance is smaller in magnitude than the object distance. The location of the image is not obvious when you look through a concave lens. In fact, since the image is smaller than the object, you may think it is farther away. But the image is closer than the object, a fact that is useful in correcting nearsightedness, as we shall see in a later section.

Table 25.3 summarizes the three types of images formed by single thin lenses. These are referred to as case 1, 2, and 3 images. Convex (converging) lenses can form either real or virtual images (cases 1 and 2, respectively), whereas concave (diverging) lenses can form only virtual images (always case 3). Real images are always inverted, but they can be either larger or smaller than the object. For example, a slide projector forms an image larger than the slide, whereas a camera makes an image smaller than the object being photographed. Virtual images are always upright and cannot be projected. Virtual images are larger than the object only in case 2, where a convex lens is used. The virtual image produced by a concave lens is always smaller than the object—a case 3 image. We can see and photograph virtual images only by using an additional lens to form a real image.

<table>
<thead>
<tr>
<th>Type</th>
<th>Formed when</th>
<th>Image type</th>
<th>( d_i )</th>
<th>( m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>( f ) positive, ( d_o &gt; f )</td>
<td>real</td>
<td>positive</td>
<td>negative</td>
</tr>
<tr>
<td>Case 2</td>
<td>( f ) positive, ( d_o &lt; f )</td>
<td>virtual</td>
<td>negative</td>
<td>positive ( m &gt; 1 )</td>
</tr>
<tr>
<td>Case 3</td>
<td>( f ) negative</td>
<td>virtual</td>
<td>negative</td>
<td>positive ( m &lt; 1 )</td>
</tr>
</tbody>
</table>

In Image Formation by Mirrors, we shall see that mirrors can form exactly the same types of images as lenses.

Take-Home Experiment: Concentrating Sunlight
Find several lenses and determine whether they are converging or diverging. In general those that are thicker near the edges are diverging and those that are thicker near the center are converging. On a bright sunny day take the converging lenses outside and try focusing the sunlight onto a piece of paper. Determine the focal lengths of the lenses. Be careful because the paper may start to burn, depending on the type of lens you have selected.

Problem-Solving Strategies for Lenses
Step 1. Examine the situation to determine that image formation by a lens is involved.
Step 2. Determine whether ray tracing, the thin lens equations, or both are to be employed. A sketch is very useful even if ray tracing is not specifically required by the problem. Write symbols and values on the sketch.

Step 3. Identify exactly what needs to be determined in the problem (identify the unknowns).

Step 4. Make a list of what is given or can be inferred from the problem as stated (identify the knowns). It is helpful to determine whether the situation involves a case 1, 2, or 3 image. While these are just names for types of images, they have certain characteristics (given in Table 25.3) that can be of great use in solving problems.

Step 5. If ray tracing is required, use the ray tracing rules listed near the beginning of this section.

Step 6. Most quantitative problems require the use of the thin lens equations. These are solved in the usual manner by substituting knowns and solving for unknowns. Several worked examples serve as guides.

Step 7. Check to see if the answer is reasonable: Does it make sense? If you have identified the type of image (case 1, 2, or 3), you should assess whether your answer is consistent with the type of image, magnification, and so on.

Misconception Alert

We do not realize that light rays are coming from every part of the object, passing through every part of the lens, and all can be used to form the final image.

We generally feel the entire lens, or mirror, is needed to form an image. Actually, half a lens will form the same, though a fainter, image.

25.7 Image Formation by Mirrors

We only have to look as far as the nearest bathroom to find an example of an image formed by a mirror. Images in flat mirrors are the same size as the object and are located behind the mirror. Like lenses, mirrors can form a variety of images. For example, dental mirrors may produce a magnified image, just as makeup mirrors do. Security mirrors in shops, on the other hand, form images that are smaller than the object. We will use the law of reflection to understand how mirrors form images, and we will find that mirror images are analogous to those formed by lenses.

Figure 25.40 helps illustrate how a flat mirror forms an image. Two rays are shown emerging from the same point, striking the mirror, and being reflected into the observer’s eye. The rays can diverge slightly, and both still get into the eye. If the rays are extrapolated backward, they seem to originate from a common point behind the mirror, locating the image. (The paths of the reflected rays into the eye are the same as if they had come directly from that point behind the mirror.) Using the law of reflection—the angle of reflection equals the angle of incidence—we can see that the image and object are the same distance from the mirror. This is a virtual image, since it cannot be projected—the rays only appear to originate from a common point behind the mirror. Obviously, if you walk behind the mirror, you cannot see the image, since the rays do not go there. But in front of the mirror, the rays behave exactly as if they had come from behind the mirror, so that is where the image is situated.

Now let us consider the focal length of a mirror—for example, the concave spherical mirrors in Figure 25.41. Rays of light that strike the surface follow the law of reflection. For a mirror that is large compared with its radius of curvature, as in Figure 25.41(a), we see that the reflected rays do not cross at the same point, and the mirror does not have a well-defined focal point. If the mirror had the shape of a parabola, the rays would all cross at a single point, and the mirror would have a well-defined focal point. But parabolic mirrors are much more expensive to make than spherical mirrors. The solution is to use a mirror that is small compared with its radius of curvature, as shown in Figure 25.41(b). (This is the mirror equivalent of the thin lens approximation.) To a very good approximation, this mirror has a well-defined focal point at F that is the focal distance \( f \) from the center of the mirror. The focal length \( f \) of a concave mirror is positive, since it is a converging mirror.
Figure 25.41 (a) Parallel rays reflected from a large spherical mirror do not all cross at a common point. (b) If a spherical mirror is small compared with its radius of curvature, parallel rays are focused to a common point. The distance of the focal point from the center of the mirror is its focal length \( f \). Since this mirror is converging, it has a positive focal length.

Just as for lenses, the shorter the focal length, the more powerful the mirror; thus, \( P = \frac{1}{f} \) for a mirror, too. A more strongly curved mirror has a shorter focal length and a greater power. Using the law of reflection and some simple trigonometry, it can be shown that the focal length is half the radius of curvature, or

\[
f = \frac{R}{2},
\]

where \( R \) is the radius of curvature of a spherical mirror. The smaller the radius of curvature, the smaller the focal length and, thus, the more powerful the mirror.

The convex mirror shown in Figure 25.42 also has a focal point. Parallel rays of light reflected from the mirror seem to originate from the point \( F \) at the focal distance \( f \) behind the mirror. The focal length and power of a convex mirror are negative, since it is a diverging mirror.

Ray tracing is as useful for mirrors as for lenses. The rules for ray tracing for mirrors are based on the illustrations just discussed:

1. A ray approaching a concave converging mirror parallel to its axis is reflected through the focal point \( F \) of the mirror on the same side. (See rays 1 and 3 in Figure 25.41(b).)
2. A ray approaching a convex diverging mirror parallel to its axis is reflected so that it seems to come from the focal point \( F \) behind the mirror. (See rays 1 and 3 in Figure 25.42.)
3. Any ray striking the center of a mirror is followed by applying the law of reflection; it makes the same angle with the axis when leaving as when approaching. (See ray 2 in Figure 25.43.)
4. A ray approaching a concave converging mirror through its focal point is reflected parallel to its axis. (The reverse of rays 1 and 3 in Figure 25.41.)
5. A ray approaching a convex diverging mirror by heading toward its focal point on the opposite side is reflected parallel to the axis. (The reverse of rays 1 and 3 in Figure 25.42.)

We will use ray tracing to illustrate how images are formed by mirrors, and we can use ray tracing quantitatively to obtain numerical information. But since we assume each mirror is small compared with its radius of curvature, we can use the thin lens equations for mirrors just as we did for lenses.

Consider the situation shown in Figure 25.43, concave spherical mirror reflection, in which an object is placed farther from a concave (converging) mirror than its focal length. That is, \( f \) is positive and \( d_o > f \), so that we may expect an image similar to the case 1 real image formed by a converging lens. Ray tracing in Figure 25.43 shows that the rays from a common point on the object all cross at a point on the same side of the mirror as the object. Thus a real image can be projected onto a screen placed at this location. The image distance is positive, and the image is inverted, so its magnification is negative. This is a case 1 image for mirrors. It differs from the case 1 image for lenses only in that the image is on the same side of the mirror as the object. It is otherwise identical.

![Figure 25.43](Image)

**Figure 25.43** A case 1 image for a mirror. An object is farther from the converging mirror than its focal length. Rays from a common point on the object are traced using the rules in the text. Ray 1 approaches parallel to the axis, ray 2 strikes the center of the mirror, and ray 3 goes through the focal point on the way toward the mirror. All three rays cross at the same point after being reflected, locating the inverted real image. Although three rays are shown, only two of the three are needed to locate the image and determine its height.

### Example 25.9 A Concave Reflector

Electric room heaters use a concave mirror to reflect infrared (IR) radiation from hot coils. Note that IR follows the same law of reflection as visible light. Given that the mirror has a radius of curvature of 50.0 cm and produces an image of the coils 3.00 m away from the mirror, where are the coils?

**Strategy and Concept**

We are given that the concave mirror projects a real image of the coils at an image distance \( d_i = 3.00 \text{ m} \). The coils are the object, and we are asked to find their location—that is, to find the object distance \( d_o \). We are also given the radius of curvature of the mirror, so that its focal length is \( f = R / 2 = 25.0 \text{ cm} \) (positive since the mirror is concave or converging). Assuming the mirror is small compared with its radius of curvature, we can use the thin lens equations, to solve this problem.

**Solution**

Since \( d_i \) and \( f \) are known, thin lens equation can be used to find \( d_o \):

\[
\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}.
\]

Rearranging to isolate \( d_o \) gives

\[
\frac{1}{d_o} = \frac{1}{f} - \frac{1}{d_i}.
\]

Entering known quantities gives a value for \( 1/d_o \):

\[
\frac{1}{d_o} = \frac{1}{0.250 \text{ m}} - \frac{1}{3.00 \text{ m}} = \frac{3.667}{\text{m}}.
\]

This must be inverted to find \( d_o \):

\[
d_o = \frac{1}{3.667} = 27.3 \text{ cm}.
\]

**Discussion**

Note that the object (the filament) is farther from the mirror than the mirror’s focal length. This is a case 1 image (\( d_o > f \) and \( f \) positive), consistent with the fact that a real image is formed. You will get the most concentrated thermal energy directly in front of the mirror and 3.00 m away from it. Generally, this is not desirable, since it could cause burns. Usually, you want the rays to emerge parallel, and this is accomplished by having the filament at the focal point of the mirror.

Note that the filament here is not much farther from the mirror than its focal length and that the image produced is considerably farther away. This is exactly analogous to a slide projector. Placing a slide only slightly farther away from the projector lens than its focal length produces an
As the object gets closer to the focal distance, the image gets farther away. In fact, as the object distance approaches the focal length, the image distance approaches infinity and the rays are sent out parallel to one another.

**Example 25.10 Solar Electric Generating System**

One of the solar technologies used today for generating electricity is a device (called a parabolic trough or concentrating collector) that concentrates the sunlight onto a blackened pipe that contains a fluid. This heated fluid is pumped to a heat exchanger, where its heat energy is transferred to another system that is used to generate steam—and so generate electricity through a conventional steam cycle. Figure 25.44 shows such a working system in southern California. Concave mirrors are used to concentrate the sunlight onto the pipe. The mirror has the approximate shape of a section of a cylinder. For the problem, assume that the mirror is exactly one-quarter of a full cylinder.

- If we wish to place the fluid-carrying pipe 40.0 cm from the concave mirror at the mirror’s focal point, what will be the radius of curvature of the mirror?
- Per meter of pipe, what will be the amount of sunlight concentrated onto the pipe, assuming the insolation (incident solar radiation) is 0.900 kW/m²?
- If the fluid-carrying pipe has a 2.00-cm diameter, what will be the temperature increase of the fluid per meter of pipe over a period of one minute? Assume all the solar radiation incident on the reflector is absorbed by the pipe, and that the fluid is mineral oil.

**Strategy**

To solve an Integrated Concept Problem we must first identify the physical principles involved. Part (a) is related to the current topic. Part (b) involves a little math, primarily geometry. Part (c) requires an understanding of heat and density.

**Solution to (a)**

To a good approximation for a concave or semi-spherical surface, the point where the parallel rays from the sun converge will be at the focal point, so \( R = 2f = 80.0 \text{ cm} \).

**Solution to (b)**

The insolation is 900 W/m². We must find the cross-sectional area \( A \) of the concave mirror, since the power delivered is 900 W/m²×A.

The mirror in this case is a quarter-section of a cylinder, so the area for a length \( L \) of the mirror is \( A = \frac{1}{4}(2\pi R)L \). The area for a length of 1.00 m is then

\[
A = \frac{\pi}{2}R(1.00 \text{ m}) = \frac{(3.14)}{2}(0.800 \text{ m})(1.00 \text{ m}) = 1.26 \text{ m}^2. \tag{25.50}
\]

The insolation on the 1.00-m length of pipe is then

\[
(9.00\times10^2 \text{ W/m}^2)(1.26 \text{ m}^2) = 1130 \text{ W}. \tag{25.51}
\]

**Solution to (c)**

The increase in temperature is given by \( Q = mc\Delta T \). The mass \( m \) of the mineral oil in the one-meter section of pipe is

\[
m = \rho V = \rho \pi \left(\frac{d}{2}\right)^2 (1.00 \text{ m}) = (8.00\times10^2 \text{ kg/m}^3)(3.14)(0.0100 \text{ m})^2 (1.00 \text{ m}) = 0.251 \text{ kg}. \tag{25.52}
\]

Therefore, the increase in temperature in one minute is

\[
\Delta T = \frac{Q}{mc} = \frac{(1130 \text{ W})(60.0 \text{ s})}{(0.251 \text{ kg})(1670 \text{ J/kg}^\circ\text{C})} = 162^\circ\text{C}. \tag{25.53}
\]

**Discussion for (c)**

An array of such pipes in the California desert can provide a thermal output of 250 MW on a sunny day, with fluids reaching temperatures as high as 400°C. We are considering only one meter of pipe here, and ignoring heat losses along the pipe.
What happens if an object is closer to a concave mirror than its focal length? This is analogous to a case 2 image for lenses \((d_o < f\) and \(f\) positive), which is a magnifier. In fact, this is how makeup mirrors act as magnifiers. Figure 25.45(a) uses ray tracing to locate the image of an object placed close to a concave mirror. Rays from a common point on the object are reflected in such a manner that they appear to be coming from behind the mirror, meaning that the image is virtual and cannot be projected. As with a magnifying glass, the image is upright and larger than the object. This is a case 2 image for mirrors and is exactly analogous to that for lenses.

![Figure 25.45](image)

Figure 25.45 (a) Case 2 images for mirrors are formed when a converging mirror has an object closer to it than its focal length. Ray 1 approaches parallel to the axis, ray 2 strikes the center of the mirror, and ray 3 approaches the mirror as if it came from the focal point. (b) A magnifying mirror showing the reflection. (credit: Mike Melrose, Flickr)

All three rays appear to originate from the same point after being reflected, locating the upright virtual image behind the mirror and showing it to be larger than the object. (b) Makeup mirrors are perhaps the most common use of a concave mirror to produce a larger, upright image.

A convex mirror is a diverging mirror \((f\) is negative) and forms only one type of image. It is a case 3 image—one that is upright and smaller than the object, just as for diverging lenses. Figure 25.46(a) uses ray tracing to illustrate the location and size of the case 3 image for mirrors. Since the image is behind the mirror, it cannot be projected and is thus a virtual image. It is also seen to be smaller than the object.
Figure 25.46 Case 3 images for mirrors are formed by any convex mirror. Ray 1 approaches parallel to the axis, ray 2 strikes the center of the mirror, and ray 3 approaches toward the focal point. All three rays appear to originate from the same point after being reflected, locating the upright virtual image behind the mirror and showing it to be smaller than the object. (b) Security mirrors are convex, producing a smaller, upright image. Because the image is smaller, a larger area is imaged compared to what would be observed for a flat mirror (and hence security is improved). (credit: Laura D’Alessandro, Flickr)

Example 25.11 Image in a Convex Mirror

A keratometer is a device used to measure the curvature of the cornea, particularly for fitting contact lenses. Light is reflected from the cornea, which acts like a convex mirror, and the keratometer measures the magnification of the image. The smaller the magnification, the smaller the radius of curvature of the cornea. If the light source is 12.0 cm from the cornea and the image’s magnification is 0.0320, what is the cornea’s radius of curvature?

**Strategy**
If we can find the focal length of the convex mirror formed by the cornea, we can find its radius of curvature (the radius of curvature is twice the focal length of a spherical mirror). We are given that the object distance is \( d_o = 12.0 \text{ cm} \) and that \( m = 0.0320 \). We first solve for the image distance \( d_i \), and then for \( f \).

**Solution**
\[ m = \frac{-d_i}{d_o} \]

Solving this expression for \( d_i \) gives
\[ d_i = -md_o \]

Entering known values yields
\[ d_i = -(0.0320)(12.0 \text{ cm}) = -0.384 \text{ cm} \]

\[ \frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \]

Substituting known values,
\[ \frac{1}{f} = \frac{1}{12.0 \text{ cm}} + \frac{1}{-0.384 \text{ cm}} = \frac{-2.52 \text{ cm}}{ \text{cm}} \]

This must be inverted to find \( f \):
\[ f = \frac{-0.400 \text{ cm}}{2.52} \]

This content is available for free at http://cnx.org/content/col11406/1.7
The radius of curvature is twice the focal length, so that
\[
R = 2 | f | = 0.800 \text{ cm.} \tag{25.59}
\]

**Discussion**

Although the focal length \( f \) of a convex mirror is defined to be negative, we take the absolute value to give us a positive value for \( R \). The radius of curvature found here is reasonable for a cornea. The distance from cornea to retina in an adult eye is about 2.0 cm. In practice, many corneas are not spherical, complicating the job of fitting contact lenses. Note that the image distance here is negative, consistent with the fact that the image is behind the mirror, where it cannot be projected. In this section’s Problems and Exercises, you will show that for a fixed object distance, the smaller the radius of curvature, the smaller the magnification.

The three types of images formed by mirrors (cases 1, 2, and 3) are exactly analogous to those formed by lenses, as summarized in the table at the end of *Image Formation by Lenses*. It is easiest to concentrate on only three types of images—then remember that concave mirrors act like convex lenses, whereas convex mirrors act like concave lenses.

**Take-Home Experiment: Concave Mirrors Close to Home**

Find a flashlight and identify the curved mirror used in it. Find another flashlight and shine the first flashlight onto the second one, which is turned off. Estimate the focal length of the mirror. You might try shining a flashlight on the curved mirror behind the headlight of a car, keeping the headlight switched off, and determine its focal length.

**Problem-Solving Strategy for Mirrors**

Step 1. Examine the situation to determine that image formation by a mirror is involved.

Step 2. Refer to the *Problem-Solving Strategies for Lenses*. The same strategies are valid for mirrors as for lenses with one qualification—use the ray tracing rules for mirrors listed earlier in this section.

**Glossary**

- **converging lens**: a convex lens in which light rays that enter it parallel to its axis converge at a single point on the opposite side
- **converging mirror**: a concave mirror in which light rays that strike it parallel to its axis converge at one or more points along the axis
- **corner reflector**: an object consisting of two mutually perpendicular reflecting surfaces, so that the light that enters is reflected back exactly parallel to the direction from which it came
- **critical angle**: incident angle that produces an angle of refraction of 90°
- **dispersion**: spreading of white light into its full spectrum of wavelengths
- **diverging lens**: a concave lens in which light rays that enter it parallel to its axis bend away (diverge) from its axis
- **diverging mirror**: a convex mirror in which light rays that strike it parallel to its axis bend away (diverge) from its axis
- **fiber optics**: transmission of light down fibers of plastic or glass, applying the principle of total internal reflection
- **focal length**: distance from the center of a lens or curved mirror to its focal point
- **focal point**: for a converging lens or mirror, the point at which converging light rays cross; for a diverging lens or mirror, the point from which diverging light rays appear to originate
- **geometric optics**: part of optics dealing with the ray aspect of light
- **index of refraction**: for a material, the ratio of the speed of light in vacuum to that in the material
- **law of reflection**: angle of reflection equals the angle of incidence
- **magnification**: ratio of image height to object height
- **mirror**: smooth surface that reflects light at specific angles, forming an image of the person or object in front of it
- **power**: inverse of focal length
- **rainbow**: dispersion of sunlight into a continuous distribution of colors according to wavelength, produced by the refraction and reflection of sunlight by water droplets in the sky
- **ray**: straight line that originates at some point
- **real image**: image that can be projected
- **refraction**: changing of a light ray’s direction when it passes through variations in matter
- **virtual image**: image that cannot be projected
Section Summary

25.1 The Ray Aspect of Light
- A straight line that originates at some point is called a ray.
- The part of optics dealing with the ray aspect of light is called geometric optics.
- Light can travel in three ways from a source to another location: (1) directly from the source through empty space; (2) through various media; (3) after being reflected from a mirror.

25.2 The Law of Reflection
- The angle of reflection equals the angle of incidence.
- A mirror has a smooth surface and reflects light at specific angles.
- Light is diffused when it reflects from a rough surface.
- Mirror images can be photographed and videotaped by instruments.

25.3 The Law of Refraction
- The changing of a light ray’s direction when it passes through variations in matter is called refraction.
- The speed of light in vacuum $c = 2.9972458 \times 10^8$ m/s $\approx 3.00 \times 10^8$ m/s.
- Index of refraction $n = \frac{c}{v}$, where $v$ is the speed of light in the material, $c$ is the speed of light in vacuum, and $n$ is the index of refraction.
- Snell’s law, the law of refraction, is stated in equation form as $n_1 \sin \theta_1 = n_2 \sin \theta_2$.

25.4 Total Internal Reflection
- The incident angle that produces an angle of refraction of $90^\circ$ is called critical angle.
- Total internal reflection is a phenomenon that occurs at the boundary between two mediums, such that if the incident angle in the first medium is greater than the critical angle, then all the light is reflected back into that medium.
- Fiber optics involves the transmission of light down fibers of plastic or glass, applying the principle of total internal reflection.
- Endoscopes are used to explore the body through various orifices or minor incisions, based on the transmission of light through optical fibers.
- Cladding prevents light from being transmitted between fibers in a bundle.
- Diamonds sparkle due to total internal reflection coupled with a large index of refraction.

25.5 Dispersion: The Rainbow and Prisms
- The spreading of white light into its full spectrum of wavelengths is called dispersion.
- Rainbows are produced by a combination of refraction and reflection and involve the dispersion of sunlight into a continuous distribution of colors.
- Dispersion produces beautiful rainbows but also causes problems in certain optical systems.

25.6 Image Formation by Lenses
- Light rays entering a converging lens parallel to its axis cross one another at a single point on the opposite side.
- For a converging lens, the focal point is the point at which converging light rays cross; for a diverging lens, the focal point is the point from which diverging light rays appear to originate.
- The distance from the center of the lens to its focal point is called the focal length $f$.
- Power $P$ of a lens is defined to be the inverse of its focal length, $P = \frac{1}{f}$.
- A lens that causes the light rays to bend away from its axis is called a diverging lens.
- Ray tracing is the technique of graphically determining the paths that light rays take.
- The image in which light rays from one point on the object actually cross at the location of the image and can be projected onto a screen, a piece of film, or the retina of an eye is called a real image.
- Thin lens equations are $\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$ and $\frac{h_i}{h_o} = -\frac{d_i}{d_o} = m$ (magnification).
- The distance of the image from the center of the lens is called image distance.
- An image that is on the same side of the lens as the object and cannot be projected on a screen is called a virtual image.

25.7 Image Formation by Mirrors
- The characteristics of an image formed by a flat mirror are: (a) The image and object are the same distance from the mirror, (b) The image is a virtual image, and (c) The image is situated behind the mirror.
- Image length is half the radius of curvature.
- A convex mirror is a diverging mirror and forms only one type of image, namely a virtual image.

Conceptual Questions

25.2 The Law of Reflection
1. Using the law of reflection, explain how powder takes the shine off of a person’s nose. What is the name of the optical effect?

25.3 The Law of Refraction
2. Diffusion by reflection from a rough surface is described in this chapter. Light can also be diffused by refraction. Describe how this occurs in a specific situation, such as light interacting with crushed ice.

3. Why is the index of refraction always greater than or equal to 1?

4. Does the fact that the light flash from lightning reaches you before its sound prove that the speed of light is extremely large or simply that it is greater than the speed of sound? Discuss how you could use this effect to get an estimate of the speed of light.

5. Will light change direction toward or away from the perpendicular when it goes from air to water? Water to glass? Glass to air?

6. Explain why an object in water always appears to be at a depth shallower than it actually is? Why do people sometimes sustain neck and spinal injuries when diving into unfamiliar ponds or waters?

7. Explain why a person's legs appear very short when wading in a pool. Justify your explanation with a ray diagram showing the path of rays from the feet to the eye of an observer who is out of the water.

8. Why is the front surface of a thermometer curved as shown?

Figure 25.47 The curved surface of the thermometer serves a purpose.

9. Suppose light were incident from air onto a material that had a negative index of refraction, say –1.3; where does the refracted light ray go?

25.4 Total Internal Reflection

10. A ring with a colorless gemstone is dropped into water. The gemstone becomes invisible when submerged. Can it be a diamond? Explain.

11. A high-quality diamond may be quite clear and colorless, transmitting all visible wavelengths with little absorption. Explain how it can sparkle with flashes of brilliant color when illuminated by white light.

12. Is it possible that total internal reflection plays a role in rainbows? Explain in terms of indices of refraction and angles, perhaps referring to Figure 25.48. Some of us have seen the formation of a double rainbow. Is it physically possible to observe a triple rainbow?

Figure 25.48 Double rainbows are not a very common observance. (credit: InvictusOU812, Flickr)

13. The most common type of mirage is an illusion that light from faraway objects is reflected by a pool of water that is not really there. Mirages are generally observed in deserts, when there is a hot layer of air near the ground. Given that the refractive index of air is lower for air at higher temperatures, explain how mirages can be formed.

25.6 Image Formation by Lenses

14. It can be argued that a flat piece of glass, such as in a window, is like a lens with an infinite focal length. If so, where does it form an image? That is, how are \(d_i\) and \(d_o\) related?

15. You can often see a reflection when looking at a sheet of glass, particularly if it is darker on the other side. Explain why you can often see a double image in such circumstances.

16. When you focus a camera, you adjust the distance of the lens from the film. If the camera lens acts like a thin lens, why can it not be a fixed distance from the film for both near and distant objects?

17. A thin lens has two focal points, one on either side, at equal distances from its center, and should behave the same for light entering from either side. Look through your eyeglasses (or those of a friend) backward and forward and comment on whether they are thin lenses.

18. Will the focal length of a lens change when it is submerged in water? Explain.
25.7 Image Formation by Mirrors

19. What are the differences between real and virtual images? How can you tell (by looking) whether an image formed by a single lens or mirror is real or virtual?

20. Can you see a virtual image? Can you photograph one? Can one be projected onto a screen with additional lenses or mirrors? Explain your responses.

21. Is it necessary to project a real image onto a screen for it to exist?

22. At what distance is an image always located—at \(d_o\), \(d_i\), or \(f\)?

23. Under what circumstances will an image be located at the focal point of a lens or mirror?

24. What is meant by a negative magnification? What is meant by a magnification that is less than 1 in magnitude?

25. Can a case 1 image be larger than the object even though its magnification is always negative? Explain.

26. Figure 25.49 shows a light bulb between two mirrors. One mirror produces a beam of light with parallel rays; the other keeps light from escaping without being put into the beam. Where is the filament of the light in relation to the focal point or radius of curvature of each mirror?

![Figure 25.49](image.png)

Figure 25.49 The two mirrors trap most of the bulb's light and form a directional beam as in a headlight.

27. Devise an arrangement of mirrors allowing you to see the back of your head. What is the minimum number of mirrors needed for this task?

28. If you wish to see your entire body in a flat mirror (from head to toe), how tall should the mirror be? Does its size depend upon your distance away from the mirror? Provide a sketch.

29. It can be argued that a flat mirror has an infinite focal length. If so, where does it form an image? That is, how are \(d_i\) and \(d_o\) related?

30. Why are diverging mirrors often used for rear-view mirrors in vehicles? What is the main disadvantage of using such a mirror compared with a flat one?
Problems & Exercises

25.1 The Ray Aspect of Light

1. Suppose a man stands in front of a mirror as shown in Figure 25.50. His eyes are 1.65 m above the floor, and the top of his head is 0.13 m higher. Find the height above the floor of the top and bottom of the smallest mirror in which he can see both the top of his head and his feet. How is this distance related to the man’s height?

25.2 The Law of Reflection

2. Show that when light reflects from two mirrors that meet each other at a right angle, the outgoing ray is parallel to the incoming ray, as illustrated in the following figure.

3. Light shows staged with lasers use moving mirrors to swing beams and create colorful effects. Show that a light ray reflected from a mirror changes direction by $2\theta$ when the mirror is rotated by an angle $\theta$.

4. A flat mirror is neither converging nor diverging. To prove this, consider two rays originating from the same point and diverging at an angle $\theta$. Show that after striking a plane mirror, the angle between their directions remains $\theta$.

25.3 The Law of Refraction

5. What is the speed of light in water? In glycerine?

6. What is the speed of light in air? In crown glass?

7. Calculate the index of refraction for a medium in which the speed of light is $2.012 \times 10^8$ m/s, and identify the most likely substance based on Table 25.1.

8. In what substance in Table 25.1 is the speed of light $2.290 \times 10^8$ m/s?

9. There was a major collision of an asteroid with the Moon in medieval times. It was described by monks at Canterbury Cathedral in England as a red glow on and around the Moon. How long after the asteroid hit the Moon, which is $3.84 \times 10^5$ km away, would the light first arrive on Earth?

10. A scuba diver training in a pool looks at his instructor as shown in Figure 25.53. What angle does the ray from the instructor’s face make with the perpendicular to the water at the point where the ray enters? The angle between the ray in the water and the perpendicular to the water is 25.0º.

11. Components of some computers communicate with each other through optical fibers having an index of refraction $n = 1.55$. What time in nanoseconds is required for a signal to travel 0.200 m through such a fiber?

12. (a) Using information in Figure 25.53, find the height of the instructor’s head above the water, noting that you will first have to calculate the angle of incidence. (b) Find the apparent depth of the diver’s head below water as seen by the instructor.
13. Suppose you have an unknown clear substance immersed in water, and you wish to identify it by finding its index of refraction. You arrange to have a beam of light enter it at an angle of $45.0^\circ$, and you observe the angle of refraction to be $40.3^\circ$. What is the index of refraction of the substance and its likely identity?

14. On the Moon’s surface, lunar astronauts placed a corner reflector, off which a laser beam is periodically reflected. The distance to the Moon is calculated from the round-trip time. What percent correction is needed to account for the delay in time due to the slowing of light in Earth’s atmosphere? Assume the distance to the Moon is precisely $3.84 \times 10^8$ m, and Earth’s atmosphere (which varies in density with altitude) is equivalent to a layer 30.0 km thick with a constant index of refraction $n = 1.000293$.

15. Suppose Figure 25.54 represents a ray of light going from air through crown glass into water, such as going into a fish tank. Calculate the amount the ray is displaced by the glass ($\Delta x$), given that the incident angle is $40.0^\circ$ and the glass is 1.00 cm thick.

16. Figure 25.54 shows a ray of light passing from one medium into a second and then a third. Show that $q_3$ is the same as it would be if the second medium were not present (provided total internal reflection does not occur).

![Figure 25.54](image)

Figure 25.54 A ray of light passes from one medium to a third by traveling through a second. The final direction is the same as if the second medium were not present, but the ray is displaced by $\Delta x$ (shown exaggerated).

17. Unreasonable Results

Suppose light travels from water to another substance, with an angle of incidence of $10.0^\circ$ and an angle of refraction of $14.9^\circ$. (a) What is the index of refraction of the other substance? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

18. Construct Your Own Problem

Consider sunlight entering the Earth’s atmosphere at sunrise and sunset—that is, at a $90^\circ$ incident angle. Taking the boundary between nearly empty space and the atmosphere to be sudden, calculate the angle of refraction for sunlight. This lengthens the time the Sun appears to be above the horizon, both at sunrise and sunset. Now construct a problem in which you determine the angle of refraction for different models of the atmosphere, such as various layers of varying density. Your instructor may wish to guide you on the level of complexity to consider and on how the index of refraction varies with air density.

19. Unreasonable Results

Light traveling from water to a gemstone strikes the surface at an angle of $80.0^\circ$ and has an angle of refraction of $15.2^\circ$. (a) What is the speed of light in the gemstone? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

20. Verify that the critical angle for light going from water to air is $48.6^\circ$, as discussed at the end of Example 25.4, regarding the critical angle for light traveling in a polystyrene (a type of plastic) pipe surrounded by air.

21. (a) At the end of Example 25.4, it was stated that the critical angle for light going from diamond to air is $24.4^\circ$. Verify this. (b) What is the critical angle for light going from zircon to air?

22. An optical fiber uses flint glass clad with crown glass. What is the critical angle?

23. At what minimum angle will you get total internal reflection of light traveling in water and reflected from ice?

24. Suppose you are using total internal reflection to make an efficient corner reflector. If there is air outside and the incident angle is $45.0^\circ$, what must be the minimum index of refraction of the material from which the reflector is made?

25. You can determine the index of refraction of a substance by determining its critical angle. (a) What is the index of refraction of a substance that has a critical angle of $68.4^\circ$ when submerged in water? What is the substance, based on Table 25.1? (b) What would the critical angle be for this substance in air?

26. A ray of light, emitted beneath the surface of an unknown liquid with air above it, undergoes total internal reflection as shown in Figure 25.55. What is the index of refraction for the liquid and its likely identification?

![Figure 25.55](image)

Figure 25.55 A light ray inside a liquid strikes the surface at the critical angle and undergoes total internal reflection.

27. A light ray entering an optical fiber surrounded by air is first refracted and then reflected as shown in Figure 25.56. Show that if the fiber is made from crown glass, any incident ray will be totally internally reflected.

![Figure 25.56](image)

Figure 25.56 A light ray enters the end of a fiber, the surface of which is perpendicular to its sides. Examine the conditions under which it may be totally internally reflected.

25.5 Dispersion: The Rainbow and Prisms

28. (a) What is the ratio of the speed of red light to violet light in diamond, based on Table 25.2? (b) What is this ratio in polystyrene? (c) Which is more dispersive?

29. A beam of white light goes from air into water at an incident angle of $75.0^\circ$. At what angles are the red (660 nm) and violet (410 nm) parts of the light refracted?

30. By how much do the critical angles for red (660 nm) and violet (410 nm) light differ in a diamond surrounded by air?
31. A narrow beam of light containing yellow (580 nm) and green (550 nm) wavelengths goes from polystyrene to air, striking the surface at a 30.0° incident angle. What is the angle between the colors when they emerge? (b) How far would they have to travel to be separated by 1.00 mm?

32. A parallel beam of light containing orange (610 nm) and violet (410 nm) wavelengths goes from fused quartz to water, striking the surface between them at a 60.0° incident angle. What is the angle between the two colors in water?

33. A ray of 610 nm light goes from air into fused quartz at an incident angle of 55.0°. At what incident angle must 470 nm light enter flint glass to have the same angle of refraction?

34. A narrow beam of light containing red (660 nm) and blue (470 nm) wavelengths travels from air through a 1.00 cm thick flat piece of crown glass and back to air again. The beam strikes at a 30.0° incident angle. (a) At what angles do the two colors emerge? (b) By what distance are the red and blue separated when they emerge?

35. A narrow beam of white light enters a prism made of crown glass at a 45.0° incident angle, as shown in Figure 25.57. At what angles, \( \theta_R \) and \( \theta_V \), do the red (660 nm) and violet (410 nm) components of the light emerge from the prism?

\[ \theta_R \] and \( \theta_V \) are the angles at which the red and violet light emerge from the prism.

36. What is the power in diopters of a camera lens that has a 50.0 mm focal length?

37. What is the focal length of a 1.75 D reading glasses found on the rack in a pharmacy?

38. You note that your prescription for new eyeglasses is –4.50 D. What will their focal length be?

39. A narrow beam of light containing orange (610 nm) and violet (410 nm) wavelengths goes from fused quartz to water, striking the surface between them at a 60.0° incident angle. What is the angle between the two colors in water?

40. A camera with a 50.0 mm focal length lens is being used to photograph a flower 75.0 cm away. How far from the lens must the film in a camera be, if the lens has a 1.75 D power?

41. A certain slide projector has a 100 mm focal length lens. How far from the lens must the film in a camera be, if the lens has a 1.75 D power?

42. A doctor examines a mole with a 15.0 cm focal length magnifying glass held 13.5 cm from the mole (a) Where is the image? (b) What is its magnification? (c) How big is the image of a 5.00 mm diameter mole?

43. A camera with a 50.0 mm focal length lens is being used to photograph a person standing 3.00 m away. (a) How far from the lens must the film be? (b) If the film is 36.0 mm high, what fraction of a 1.75 m tall person will fit on it? (c) Discuss how reasonable this seems, based on your experience in taking or posing for photographs.

44. A camera lens used for taking close-up photographs has a focal length of 22.0 mm. The farthest it can be placed from the film is 33.0 mm. (a) What is the closest object that can be photographed? (b) What is the magnification of this closest object?

45. A camera with a 50.0 mm focal length lens is being used to photograph a flower 75.0 cm away. How far from the lens must the film in a camera be, if the lens has a 1.75 D power?

46. Suppose your 50.0 mm focal length camera lens is 51.0 mm away from the film in the camera. (a) How far away is an object that is in focus? (b) What is the height of the object if its image is 2.00 cm high?

47. A camera lens used for taking close-up photographs has a focal length of 22.0 mm. The farthest it can be placed from the film is 33.0 mm. What is the closest object that can be photographed? (b) What is the magnification of this closest object?

48. A camera lens used for taking close-up photographs has a focal length of 22.0 mm. The farthest it can be placed from the film is 33.0 mm. (a) What is the closest object that can be photographed? (b) What is the magnification of this closest object?

49. Suppose a 200 mm focal length telephoto lens is being used to photograph mountains 10.0 km away. (a) Where is the image? (b) What is the height of the image of a 1000 m high cliff on one of the mountains?

50. A camera with a 100 mm focal length lens is used to photograph the sun and moon. What is the height of the image of the sun on the film, given the sun is 1.40 \( \times 10^6 \) km in diameter and is 1.50 \( \times 10^8 \) km away?

51. A camera lens used for taking close-up photographs has a focal length of 22.0 mm. The farthest it can be placed from the film is 33.0 mm. (a) What is the closest object that can be photographed? (b) What is the magnification of this closest object?

52. A camera lens used for taking close-up photographs has a focal length of 22.0 mm. The farthest it can be placed from the film is 33.0 mm. (a) What is the closest object that can be photographed? (b) What is the magnification of this closest object?

53. What is the focal length of a makeup mirror that has a power of 1.50 D?

54. Some telephoto cameras use a mirror rather than a lens. What radius of curvature mirror is needed to replace a 800 mm focal length telephoto lens?

55. A camera lens used for taking close-up photographs has a focal length of 22.0 mm. The farthest it can be placed from the film is 33.0 mm. (a) What is the closest object that can be photographed? (b) What is the magnification of this closest object?
61. Show that for a flat mirror \( h_i = h_o \), knowing that the image is a distance behind the mirror equal in magnitude to the distance of the object from the mirror.

62. Use the law of reflection to prove that the focal length of a mirror is half its radius of curvature. That is, prove that \( f = R/2 \). Note this is true for a spherical mirror only if its diameter is small compared with its radius of curvature.

63. Referring to the electric room heater considered in the first example in this section, calculate the intensity of IR radiation in \( \text{W/m}^2 \) projected by the concave mirror on a person 3.00 m away. Assume that the heating element radiates 1500 W and has an area of 100 cm\(^2\), and that half of the radiated power is reflected and focused by the mirror.

64. Consider a 250-W heat lamp fixed to the ceiling in a bathroom. If the filament in one light burns out then the remaining three still work. Construct a problem in which you determine the resistance of each filament in order to obtain a certain intensity projected on the bathroom floor. The ceiling is 3.0 m high. The problem will need to involve concave mirrors behind the filaments. Your instructor may wish to guide you on the level of complexity to consider in the electrical components.
Learning Objectives

26.1. Physics of the Eye
- Explain the image formation by the eye.
- Explain why peripheral images lack detail and color.
- Define refractive indices.
- Analyze the accommodation of the eye for distant and near vision.

26.2. Vision Correction
- Identify and discuss common vision defects.
- Explain nearsightedness and farsightedness corrections.
- Explain laser vision correction.

26.3. Color and Color Vision
- Explain the simple theory of color vision.
- Outline the coloring properties of light sources.
- Describe the retinex theory of color vision.

26.4. Microscopes
- Investigate different types of microscopes.
- Learn how image is formed in a compound microscope.

26.5. Telescopes
- Outline the invention of a telescope.
- Describe the working of a telescope.

26.6. Aberrations
- Describe optical aberration.

Introduction to Vision and Optical Instruments

Explore how the image on the computer screen is formed. How is the image formation on the computer screen different from the image formation in your eye as you look down the microscope? How can videos of living cell processes be taken for viewing later on, and by many different people?
Seeing faces and objects we love and cherish is a delight—one’s favorite teddy bear, a picture on the wall, or the sun rising over the mountains. Intricate images help us understand nature and are invaluable for developing techniques and technologies in order to improve the quality of life. The image of a red blood cell that almost fills the cross-sectional area of a tiny capillary makes us wonder how blood makes it through and not get stuck. We are able to see bacteria and viruses and understand their structure. It is the knowledge of physics that provides fundamental understanding and models required to develop new techniques and instruments. Therefore, physics is called an enabling science—a science that enables development and advancement in other areas. It is through optics and imaging that physics enables advancement in major areas of biosciences. This chapter illustrates the enabling nature of physics through an understanding of how a human eye is able to see and how we are able to use optical instruments to see beyond what is possible with the naked eye. It is convenient to categorize these instruments on the basis of geometric optics (see Geometric Optics) and wave optics (see Wave Optics).

### 26.1 Physics of the Eye

The eye is perhaps the most interesting of all optical instruments. The eye is remarkable in how it forms images and in the richness of detail and color it can detect. However, our eyes commonly need some correction, to reach what is called “normal” vision, but should be called ideal rather than normal. Image formation by our eyes and common vision correction are easy to analyze with the optics discussed in Geometric Optics.

Figure 26.2 shows the basic anatomy of the eye. The cornea and lens form a system that, to a good approximation, acts as a single thin lens. For clear vision, a real image must be projected onto the light-sensitive retina, which lies at a fixed distance from the lens. The lens of the eye adjusts its power to produce an image on the retina for objects at different distances. The center of the image falls on the fovea, which has the greatest density of light receptors and the greatest acuity (sharpness) in the visual field. The variable opening (or pupil) of the eye along with chemical adaptation allows the eye to detect light intensities from the lowest observable to $10^{10}$ times greater (without damage). This is an incredible range of detection.

Our eyes perform a vast number of functions, such as sense direction, movement, sophisticated colors, and distance. Processing of visual nerve impulses begins with interconnections in the retina and continues in the brain. The optic nerve conveys signals received by the eye to the brain.

Figure 26.3 shows image formation by the cornea and lens of the eye. The rays bend according to the refractive indices provided in Table 26.1. The cornea provides about two-thirds of the power needed to produce an image on the retina. The cornea and lens can be treated as a single thin lens, even though the light rays pass through several layers of material (such as cornea, aqueous humor, several layers in the lens, and vitreous humor), changing direction at each interface. The image formed is much like the one produced by a single convex lens. This is a case 1 image. Images formed in the eye are inverted but the brain inverts them once more to make them seem upright.

<table>
<thead>
<tr>
<th>Material</th>
<th>Index of Refraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>1.33</td>
</tr>
<tr>
<td>Air</td>
<td>1.0</td>
</tr>
<tr>
<td>Cornea</td>
<td>1.38</td>
</tr>
<tr>
<td>Aqueous humor</td>
<td>1.34</td>
</tr>
<tr>
<td>Lens (average)</td>
<td>1.41 (varies throughout the lens, greatest in center)</td>
</tr>
<tr>
<td>Vitreous humor</td>
<td>1.34</td>
</tr>
</tbody>
</table>
Figure 26.3 An image is formed on the retina with light rays converging most at the cornea and upon entering and exiting the lens. Rays from the top and bottom of the object are traced and produce an inverted real image on the retina. The distance to the object is drawn smaller than scale.

As noted, the image must fall precisely on the retina to produce clear vision — that is, the image distance $d_i$ must equal the lens-to-retina distance. Because the lens-to-retina distance does not change, the image distance $d_i$ must be the same for objects at all distances. The eye manages this by varying the power (and focal length) of the lens to accommodate for objects at various distances. The process of adjusting the eye’s focal length is called accommodation. A person with normal (ideal) vision can see objects clearly at distances ranging from 25 cm to essentially infinity. However, although the near point (the shortest distance at which a sharp focus can be obtained) increases with age (becoming meters for some older people), we will consider it to be 25 cm in our treatment here.

Figure 26.4 shows the accommodation of the eye for distant and near vision. Since light rays from a nearby object can diverge and still enter the eye, the lens must be more converging (more powerful) for close vision than for distant vision. To be more converging, the lens is made thicker by the action of the ciliary muscle surrounding it. The eye is most relaxed when viewing distant objects, one reason that microscopes and telescopes are designed to produce distant images. Vision of very distant objects is called totally relaxed, while close vision is termed accommodated, with the closest vision being fully accommodated.

Figure 26.4 Relaxed and accommodated vision for distant and close objects. (a) Light rays from the same point on a distant object must be nearly parallel while entering the eye and more easily converge to produce an image on the retina. (b) Light rays from a nearby object can diverge more and still enter the eye. A more powerful lens is needed to converge them on the retina than if they were parallel.
We will use the thin lens equations to examine image formation by the eye quantitatively. First, note the power of a lens is given as $P = 1/f$, so we rewrite the thin lens equations as

$$P = \frac{1}{d_o} + \frac{1}{d_i}$$  \hspace{1cm} (26.1)

and

$$\frac{h_i}{h_o} = -\frac{d_i}{d_o} = m.$$  \hspace{1cm} (26.2)

We understand that $d_i$ must equal the lens-to-retina distance to obtain clear vision, and that normal vision is possible for objects at distances $d_o = 25 \text{ cm}$ to infinity.

**Take-Home Experiment: The Pupil**

Look at the central transparent area of someone’s eye, the pupil, in normal room light. Estimate the diameter of the pupil. Now turn off the lights and darken the room. After a few minutes turn on the lights and promptly estimate the diameter of the pupil. What happens to the pupil as the eye adjusts to the room light? Explain your observations.

The eye can detect an impressive amount of detail, considering how small the image is on the retina. To get some idea of how small the image can be, consider the following example.

**Example 26.1 Size of Image on Retina**

What is the size of the image on the retina of a $1.20 \times 10^{-2}$ cm diameter human hair, held at arm’s length (60.0 cm) away? Take the lens-to-retina distance to be 2.00 cm.

**Strategy**

We want to find the height of the image $h_i$, given the height of the object is $h_o = 1.20 \times 10^{-2}$ cm. We also know that the object is 60.0 cm away, so that $d_o = 60.0 \text{ cm}$. For clear vision, the image distance must equal the lens-to-retina distance, and so $d_i = 2.00 \text{ cm}$. The equation $\frac{h_i}{h_o} = -\frac{d_i}{d_o} = m$ can be used to find $h_i$ with the known information.

**Solution**

The only unknown variable in the equation $\frac{h_i}{h_o} = -\frac{d_i}{d_o} = m$ is $h_i$:

$$\frac{h_i}{h_o} = -\frac{d_i}{d_o}.$$  \hspace{1cm} (26.3)

Rearranging to isolate $h_i$ yields

$$h_i = -h_o \cdot \frac{d_i}{d_o}.$$  \hspace{1cm} (26.4)

Substituting the known values gives

$$h_i = -(1.20 \times 10^{-2} \text{ cm}) \frac{2.00 \text{ cm}}{60.0 \text{ cm}} = -4.00 \times 10^{-4} \text{ cm}.$$  \hspace{1cm} (26.5)

**Discussion**

This truly small image is not the smallest discernible—that is, the limit to visual acuity is even smaller than this. Limitations on visual acuity have to do with the wave properties of light and will be discussed in the next chapter. Some limitation is also due to the inherent anatomy of the eye and processing that occurs in our brain.

**Example 26.2 Power Range of the Eye**

Calculate the power of the eye when viewing objects at the greatest and smallest distances possible with normal vision, assuming a lens-to-retina distance of 2.00 cm (a typical value).

**Strategy**


For clear vision, the image must be on the retina, and so $d_i = 2.00 \text{ cm}$ here. For distant vision, $d_o \approx \infty$, and for close vision, $d_o = 25.0 \text{ cm}$, as discussed earlier. The equation $P = \frac{1}{d_o} + \frac{1}{d_i}$ as written just above, can be used directly to solve for $P$ in both cases, since we know $d_i$ and $d_o$. Power has units of diopters, where 1 D = 1/m, and so we should express all distances in meters.

Solution
For distant vision,
\[ P = \frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{\infty} + \frac{1}{0.0200 \text{ m}}. \]
Since $1/\infty = 0$, this gives
\[ P = 0 + 50.0/\text{m} = 50.0 \text{ D} \text{ (distant vision)}. \]
Now, for close vision,
\[ P = \frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{0.250 \text{ m}} + \frac{1}{0.0200 \text{ m}} \]
\[ = \frac{4.00}{\text{m}} + \frac{50.0}{\text{m}} = 4.00 \text{ D} + 50.0 \text{ D} \]
\[ = 54.0 \text{ D} \text{ (close vision)}. \]

Discussion
For an eye with this typical 2.00 cm lens-to-retina distance, the power of the eye ranges from 50.0 D (for distant totally relaxed vision) to 54.0 D (for close fully accommodated vision), which is an 8% increase. This increase in power for close vision is consistent with the preceding discussion and the ray tracing in Figure 26.4. An 8% ability to accommodate is considered normal but is typical for people who are about 40 years old. Younger people have greater accommodation ability, whereas older people gradually lose the ability to accommodate. When an optometrist identifies accommodation as a problem in elder people, it is most likely due to stiffening of the lens. The lens of the eye changes with age in ways that tend to preserve the ability to see distant objects clearly but do not allow the eye to accommodate for close vision, a condition called presbyopia (literally, elder eye). To correct this vision defect, we place a converging, positive power lens in front of the eye, such as found in reading glasses. Commonly available reading glasses are rated by their power in diopters, typically ranging from 1.0 to 3.5 D.

26.2 Vision Correction
The need for some type of vision correction is very common. Common vision defects are easy to understand, and some are simple to correct. Figure 26.5 illustrates two common vision defects. Nearsightedness, or myopia, is the inability to see distant objects clearly while close objects are clear. The eye overconverges the nearly parallel rays from a distant object, and the rays cross in front of the retina. More divergent rays from a close object are converged on the retina for a clear image. The distance to the farthest object that can be seen clearly is called the far point of the eye (normally infinity). Farsightedness, or hyperopia, is the inability to see close objects clearly while distant objects may be clear. A farsighted eye does not converge sufficient rays from a close object to make the rays meet on the retina. Less diverging rays from a distant object can be converged for a clear image. The distance to the closest object that can be seen clearly is called the near point of the eye (normally 25 cm).

![Figure 26.5](image)

Figure 26.5 (a) The nearsighted (myopic) eye converges rays from a distant object in front of the retina; thus, they are diverging when they strike the retina, producing a blurry image. This can be caused by the lens of the eye being too powerful or the length of the eye being too great. (b) The farsighted (hyperopic) eye is unable to converge the rays from a close object by the time they strike the retina, producing blurry close vision. This can be caused by insufficient power in the lens or by the eye being too short.

Since the nearsighted eye over converges light rays, the correction for nearsightedness is to place a diverging spectacle lens in front of the eye. This reduces the power of an eye that is too powerful. Another way of thinking about this is that a diverging spectacle lens produces a case 3 image,
which is closer to the eye than the object (see Figure 26.6). To determine the spectacle power needed for correction, you must know the person's far point—that is, you must know the greatest distance at which the person can see clearly. Then the image produced by a spectacle lens must be at this distance or closer for the nearsighted person to be able to see it clearly. It is worth noting that wearing glasses does not change the eye in any way. The eyeglass lens is simply used to create an image of the object at a distance where the nearsighted person can see it clearly. Whereas someone not wearing glasses can see clearly objects that fall between their near point and their far point, someone wearing glasses can see images that fall between their near point and their far point.

![Diagram of eye and image](Figure 26.6 Correction of nearsightedness requires a diverging lens that compensates for the overconvergence by the eye. The diverging lens produces an image closer to the eye than the object, so that the nearsighted person can see it clearly.)

**Example 26.3 Correcting Nearsightedness**

What power of spectacle lens is needed to correct the vision of a nearsighted person whose far point is 30.0 cm? Assume the spectacle (corrective) lens is held 1.50 cm away from the eye by eyeglass frames.

**Strategy**

You want this nearsighted person to be able to see very distant objects clearly. That means the spectacle lens must produce an image 30.0 cm from the eye for an object very far away. An image 30.0 cm from the eye will be 28.5 cm to the left of the spectacle lens (see Figure 26.6). Therefore, we must get \( d_i = -28.5 \text{ cm} \) when \( d_o \approx \infty \). The image distance is negative, because it is on the same side of the spectacle as the object.

**Solution**

Since \( d_i \) and \( d_o \) are known, the power of the spectacle lens can be found using 

\[
P = \frac{1}{d_o} + \frac{1}{d_i}
\]

as written earlier:

\[
P = \frac{1}{\infty} + \frac{1}{-0.285 \text{ m}}.
\]

Since \( 1/ \infty = 0 \), we obtain:

\[
P = 0 - 3.51 / \text{m} = -3.51 \text{ D}.
\]

**Discussion**

The negative power indicates a diverging (or concave) lens, as expected. The spectacle produces a case 3 image closer to the eye, where the person can see it. If you examine eyeglasses for nearsighted people, you will find the lenses are thinnest in the center. Additionally, if you examine a prescription for eyeglasses for nearsighted people, you will find that the prescribed power is negative and given in units of diopters.

Since the farsighted eye underconverges light rays, the correction for farsightedness is to place a converging spectacle lens in front of the eye. This increases the power of an eye that is too weak. Another way of thinking about this is that a converging spectacle lens produces a case 2 image, which is farther from the eye than the object (see Figure 26.7). To determine the spectacle power needed for correction, you must know the person's near point—that is, you must know the smallest distance at which the person can see clearly. Then the image produced by a spectacle lens must be at this distance or farther for the farsighted person to be able to see it clearly.
Example 26.4 Correcting Farsightedness

What power of spectacle lens is needed to allow a farsighted person, whose near point is 1.00 m, to see an object clearly that is 25.0 cm away? Assume the spectacle (corrective) lens is held 1.50 cm away from the eye by eyeglass frames.

Strategy

When an object is held 25.0 cm from the person’s eyes, the spectacle lens must produce an image 1.00 m away (the near point). An image 1.00 m from the eye will be 98.5 cm to the left of the spectacle lens because the spectacle lens is 1.50 cm from the eye (see Figure 26.7). Therefore, \( d_i = -98.5 \) cm. The image distance is negative, because it is on the same side of the spectacle as the object. The object is 23.5 cm to the left of the spectacle, so that \( d_o = 23.5 \) cm.

Solution

Since \( d_i \) and \( d_o \) are known, the power of the spectacle lens can be found using \( P = \frac{1}{d_o} + \frac{1}{d_i} \):

\[
P = \frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{0.235 \text{ m}} + \frac{1}{-0.985 \text{ m}} = 4.26 \text{ D} - 1.02 \text{ D} = 3.24 \text{ D}.
\]

Discussion

The positive power indicates a converging (convex) lens, as expected. The convex spectacle produces a case 2 image farther from the eye, where the person can see it. If you examine eyeglasses of farsighted people, you will find the lenses to be thickest in the center. In addition, a prescription of eyeglasses for farsighted people has a prescribed power that is positive.

Another common vision defect is astigmatism, an unevenness or asymmetry in the focus of the eye. For example, rays passing through a vertical region of the eye may focus closer than rays passing through a horizontal region, resulting in the image appearing elongated. This is mostly due to irregularities in the shape of the cornea but can also be due to lens irregularities or unevenness in the retina. Because of these irregularities, different parts of the lens system produce images at different locations. The eye-brain system can compensate for some of these irregularities, but they generally manifest themselves as less distinct vision or sharper images along certain axes. Figure 26.8 shows a chart used to detect astigmatism. Astigmatism can be at least partially corrected with a spectacle having the opposite irregularity of the eye. If an eyeglass prescription has a cylindrical correction, it is there to correct astigmatism. The normal corrections for short- or farsightedness are spherical corrections, uniform along all axes.
Contact lenses have advantages over glasses beyond their cosmetic aspects. One problem with glasses is that as the eye moves, it is not at a fixed distance from the spectacle lens. Contacts rest on and move with the eye, eliminating this problem. Because contacts cover a significant portion of the cornea, they provide superior peripheral vision compared with eyeglasses. Contacts also correct some corneal astigmatism caused by surface irregularities. The tear layer between the smooth contact and the cornea fills in the irregularities. Since the index of refraction of the tear layer and the cornea are very similar, you now have a regular optical surface in place of an irregular one. If the curvature of a contact lens is not the same as the cornea (as may be necessary with some individuals to obtain a comfortable fit), the tear layer between the contact and cornea acts as a lens. If the tear layer is thinner in the center than at the edges, it has a negative power, for example. Skilled optometrists will adjust the power of the contact to compensate.

Laser vision correction has progressed rapidly in the last few years. It is the latest and by far the most successful in a series of procedures that correct vision by reshaping the cornea. As noted at the beginning of this section, the cornea accounts for about two-thirds of the power of the eye. Thus, small adjustments of its curvature have the same effect as putting a lens in front of the eye. To a reasonable approximation, the power of multiple lenses placed close together equals the sum of their powers. For example, a concave spectacle lens (for nearsightedness) having \( P = -3.00 \text{ D} \) has the same effect on vision as reducing the power of the eye itself by 3.00 D. So to correct the eye for nearsightedness, the cornea is flattened to reduce its power. Similarly, to correct for farsightedness, the curvature of the cornea is enhanced to increase the power of the eye—the same effect as the positive power spectacle lens used for farsightedness. Laser vision correction uses high intensity electromagnetic radiation to ablate (to remove material from the surface) and reshape the corneal surfaces.

Today, the most commonly used laser vision correction procedure is Laser in situ Keratomileusis (LASIK). The top layer of the cornea is surgically peeled back and the underlying tissue ablated by multiple bursts of finely controlled ultraviolet radiation produced by an excimer laser. Lasers are used because they not only produce well-focused intense light, but they also emit very pure wavelength electromagnetic radiation that can be controlled more accurately than mixed wavelength light. The 193 nm wavelength UV commonly used is extremely and strongly absorbed by corneal tissue, allowing precise evaporation of very thin layers. A computer controlled program applies more bursts, usually at a rate of 10 per second, to the areas that require deeper removal. Typically a spot less than 1 mm in diameter and about 0.3 \( \mu \text{m} \) in thickness is removed by each burst.

Nearsightedness, farsightedness, and astigmatism can be corrected with an accuracy that produces normal distant vision in more than 90% of the patients, in many cases right away. The corneal flap is replaced; healing takes place rapidly and is nearly painless. More than 1 million Americans per year undergo LASIK (see Figure 26.9).

![Figure 26.9](link) Laser vision correction is being performed using the LASIK procedure. Reshaping of the cornea by laser ablation is based on a careful assessment of the patient’s vision and is computer controlled. The upper corneal layer is temporarily peeled back and minimally disturbed in LASIK, providing for more rapid and less painful healing of the less sensitive tissues below. (credit: U.S. Navy photo by Mass Communication Specialist 1st Class Brien Aho)

### 26.3 Color and Color Vision

The gift of vision is made richer by the existence of color. Objects and lights abound with thousands of hues that stimulate our eyes, brains, and emotions. Two basic questions are addressed in this brief treatment—what does color mean in scientific terms, and how do we, as humans, perceive it?

**Simple Theory of Color Vision**

We have already noted that color is associated with the wavelength of visible electromagnetic radiation. When our eyes receive pure-wavelength light, we tend to see only a few colors. Six of these (most often listed) are red, orange, yellow, green, blue, and violet. These are the rainbow of colors...
produced when white light is dispersed according to different wavelengths. There are thousands of other hues that we can perceive. These include brown, teal, gold, pink, and white. One simple theory of color vision implies that all these hues are our eye’s response to different combinations of wavelengths. This is true to an extent, but we find that color perception is even subtler than our eye’s response for various wavelengths of light.

The two major types of light-sensing cells (photoreceptors) in the retina are rods and cones. Rods are more sensitive than cones by a factor of about 1000 and are solely responsible for peripheral vision as well as vision in very dark environments. They are also important for motion detection. There are about 120 million rods in the human retina. Rods do not yield color information. You may notice that you lose color vision when it is very dark, but you retain the ability to discern grey scales.

Cones are most concentrated in the fovea, the central region of the retina. There are no rods here. The fovea is at the center of the macula, a 5 mm diameter region responsible for our central vision. The cones work best in bright light and are responsible for high resolution vision. There are about 6 million cones in the human retina. There are three types of cones, and each type is sensitive to different ranges of wavelengths, as illustrated in Figure 26.10. A simplified theory of color vision is that there are three primary colors corresponding to the three types of cones. The thousands of other hues that we can distinguish among are created by various combinations of stimulations of the three types of cones. Color television uses a three-color system in which the screen is covered with equal numbers of red, green, and blue phosphor dots. The broad range of hues a viewer sees is produced by various combinations of these three colors. For example, you will perceive yellow when red and green are illuminated with the correct ratio of intensities. White may be sensed when all three are illuminated. Then, it would seem that all hues can be produced by adding three primary colors in various proportions. But there is an indication that color vision is more sophisticated. There is no unique set of three primary colors. Another set that works is yellow, green, and blue. A further indication of the need for a more complex theory of color vision is that various different combinations can produce the same hue. Yellow can be sensed with yellow light, or with a combination of red and green, and also with white light from which violet has been removed. The three-primary-colors aspect of color vision is well established; more sophisticated theories expand on it rather than deny it.

Figure 26.10 The image shows the relative sensitivity of the three types of cones, which are named according to wavelengths of greatest sensitivity. Rods are about 1000 times more sensitive, and their curve peaks at about 500 nm. Evidence for the three types of cones comes from direct measurements in animal and human eyes and testing of color blind people.

Consider why various objects display color—that is, why are feathers blue and red in a crimson rosella? The true color of an object is defined by its absorptive or reflective characteristics. Figure 26.11 shows white light falling on three different objects, one pure blue, one pure red, and one black, as well as pure red light falling on a white object. Other hues are created by more complex absorption characteristics. Pink, for example on a galah cockatoo, can be due to weak absorption of all colors except red. An object can appear a different color under non-white illumination. For example, a pure blue object illuminated with pure red light will appear black, because it absorbs all the red light falling on it. But, the true color of the object is blue, which is independent of illumination.

Figure 26.11 Absorption characteristics determine the true color of an object. Here, three objects are illuminated by white light, and one by pure red light. White is the equal mixture of all visible wavelengths; black is the absence of light.
Similarly, light sources have colors that are defined by the wavelengths they produce. A helium-neon laser emits pure red light. In fact, the phrase “pure red light” is defined by having a sharp constrained spectrum, a characteristic of laser light. The Sun produces a broad yellowish spectrum, fluorescent lights emit bluish-white light, and incandescent lights emit reddish-white hues as seen in Figure 26.12. As you would expect, you sense these colors when viewing the light source directly or when illuminating a white object with them. All of this fits neatly into the simplified theory that a combination of wavelengths produces various hues.

Take-Home Experiment: Exploring Color Addition

This activity is best done with plastic sheets of different colors as they allow more light to pass through to our eyes. However, thin sheets of paper and fabric can also be used. Overlay different colors of the material and hold them up to a white light. Using the theory described above, explain the colors you observe. You could also try mixing different crayon colors.

Figure 26.12 Emission spectra for various light sources are shown. Curve A is average sunlight at Earth’s surface, curve B is light from a fluorescent lamp, and curve C is the output of an incandescent light. The spike for a helium-neon laser (curve D) is due to its pure wavelength emission. The spikes in the fluorescent output are due to atomic spectra—a topic that will be explored later.

Color Constancy and a Modified Theory of Color Vision

The eye-brain color-sensing system can, by comparing various objects in its view, perceive the true color of an object under varying lighting conditions—an ability that is called color constancy. We can sense that a white tablecloth, for example, is white whether it is illuminated by sunlight, fluorescent light, or candlelight. The wavelengths entering the eye are quite different in each case, as the graphs in Figure 26.12 imply, but our color vision can detect the true color by comparing the tablecloth with its surroundings.

Theories that take color constancy into account are based on a large body of anatomical evidence as well as perceptual studies. There are nerve connections among the light receptors on the retina, and there are far fewer nerve connections to the brain than there are rods and cones. This means that there is signal processing in the eye before information is sent to the brain. For example, the eye makes comparisons between adjacent light receptors and is very sensitive to edges as seen in Figure 26.13. Rather than responding simply to the light entering the eye, which is uniform in the various rectangles in this figure, the eye responds to the edges and senses false darkness variations.

Figure 26.13 The importance of edges is shown. Although the grey strips are uniformly shaded, as indicated by the graph immediately below them, they do not appear uniform at all. Instead, they are perceived darker on the dark side and lighter on the light side of the edge, as shown in the bottom graph. This is due to nerve impulse processing in the eye.
One theory that takes various factors into account was advanced by Edwin Land (1909 – 1991), the creative founder of the Polaroid Corporation. Land proposed, based partly on his many elegant experiments, that the three types of cones are organized into systems called retinexes. Each retinex forms an image that is compared with the others, and the eye-brain system thus can compare a candle-illuminated white table cloth with its generally reddish surroundings and determine that it is actually white. This retinex theory of color vision is an example of modified theories of color vision that attempt to account for its subtleties. One striking experiment performed by Land demonstrates that some type of image comparison may produce color vision. Two pictures are taken of a scene on black-and-white film, one using a red filter, the other a blue filter. Resulting black-and-white slides are then projected and superimposed on a screen, producing a black-and-white image, as expected. Then a red filter is placed in front of the slide taken with a red filter, and the images are again superimposed on a screen. You would expect an image in various shades of pink, but instead, the image appears to humans in full color with all the hues of the original scene. This implies that color vision can be induced by comparison of the black-and-white and red images. Color vision is not completely understood or explained, and the retinex theory is not totally accepted. It is apparent that color vision is much subtler than what a first look might imply.

PhET Explorations: Color Vision

Make a whole rainbow by mixing red, green, and blue light. Change the wavelength of a monochromatic beam or filter white light. View the light as a solid beam, or see the individual photons.

PhET Interactive Simulation

26.4 Microscopes

Although the eye is marvelous in its ability to see objects large and small, it obviously has limitations to the smallest details it can detect. Human desire to see beyond what is possible with the naked eye led to the use of optical instruments. In this section we will examine microscopes, instruments for enlarging the detail that we cannot see with the unaided eye. The microscope is a multiple-element system having more than a single lens or mirror. (See Figure 26.15) A microscope can be made from two convex lenses. The image formed by the first element becomes the object for the second element. The second element forms its own image, which is the object for the third element, and so on. Ray tracing helps to visualize the image formed. If the device is composed of thin lenses and mirrors that obey the thin lens equations, then it is not difficult to describe their behavior numerically.

Microscopes were first developed in the early 1600s by eyeglass makers in The Netherlands and Denmark. The simplest compound microscope is constructed from two convex lenses as shown schematically in Figure 26.16. The first lens is called the objective lens, and has typical magnification values from 5× to 100×. In standard microscopes, the objectives are mounted such that when you switch between objectives, the sample remains in focus. Objectives arranged in this way are described as parfocal. The second, the eyepiece, also referred to as the ocular, has several lenses which slide inside a cylindrical barrel. The focusing ability is provided by the movement of both the objective lens and the eyepiece. The purpose of a microscope is to magnify small objects, and both lenses contribute to the final magnification. Additionally, the final enlarged image is produced in a location far enough from the observer to be easily viewed, since the eye cannot focus on objects or images that are too close.
To see how the microscope in Figure 26.16 forms an image, we consider its two lenses in succession. The object is slightly farther away from the objective lens than its focal length $f_o$, producing a case 1 image that is larger than the object. This first image is the object for the second lens, or eyepiece. The eyepiece is intentionally located so it can further magnify the image. The eyepiece is placed so that the first image is closer to it than its focal length $f_e$. Thus the eyepiece acts as a magnifying glass, and the final image is made even larger. The final image remains inverted, but it is farther from the observer, making it easy to view (the eye is most relaxed when viewing distant objects and normally cannot focus closer than 25 cm). Since each lens produces a magnification that multiplies the height of the image, it is apparent that the overall magnification $m$ is the product of the individual magnifications:

$$m = m_o m_e,$$

where $m_o$ is the magnification of the objective and $m_e$ is the magnification of the eyepiece. This equation can be generalized for any combination of thin lenses and mirrors that obey the thin lens equations.

**Overall Magnification**

The overall magnification of a multiple-element system is the product of the individual magnifications of its elements.

**Example 26.5 Microscope Magnification**

Calculate the magnification of an object placed 6.20 mm from a compound microscope that has a 6.00 mm focal length objective and a 50.0 mm focal length eyepiece. The objective and eyepiece are separated by 23.0 cm.

**Strategy and Concept**

This situation is similar to that shown in Figure 26.16. To find the overall magnification, we must find the magnification of the objective, then the magnification of the eyepiece. This involves using the thin lens equation.

**Solution**

The magnification of the objective lens is given as

$$m_o = -\frac{d_i}{d_o}$$

where $d_o$ and $d_i$ are the object and image distances, respectively, for the objective lens as labeled in Figure 26.16. The object distance is given to be $d_o = 6.20$ mm, but the image distance $d_i$ is not known. Isolating $d_i$, we have

$$\frac{1}{d_i} = \frac{1}{f_o} - \frac{1}{d_o},$$

where $f_o$ is the focal length of the objective lens. Substituting known values gives

$$\frac{1}{d_i} = \frac{1}{6.00 \text{ mm}} - \frac{1}{6.20 \text{ mm}} = 0.00538 \text{ mm}^{-1}.$$  

We invert this to find $d_i$:

$$d_i = 186 \text{ mm}. $$
Substituting this into the expression for \( m_o \) gives

\[
m_o = \frac{d_i}{d_o} = \frac{-186 \text{ mm}}{6.20 \text{ mm}} = -30.0.
\]  

(26.17)

Now we must find the magnification of the eyepiece, which is given by

\[
m_e = \frac{d_{i}'}{d_{o}'}.
\]

(26.18)

where \( d_{i}' \) and \( d_{o}' \) are the image and object distances for the eyepiece (see Figure 26.16). The object distance is the distance of the first image from the eyepiece. Since the first image is 186 mm to the right of the objective and the eyepiece is 230 mm to the right of the objective, the object distance is \( d_{o}' = 230 \text{ mm} - 186 \text{ mm} = 44.0 \text{ mm} \). This places the first image closer to the eyepiece than its focal length, so that the eyepiece will form a case 2 image as shown in the figure. We still need to find the location of the final image \( d_{i}' \) in order to find the magnification. This is done as before to obtain a value for \( 1/d_{i}' \):

\[
\frac{1}{d_{i}'} = \frac{1}{f_{e}} - \frac{1}{d_{o}'} = \frac{1}{50.0 \text{ mm}} - \frac{1}{44.0 \text{ mm}} = -0.00273 \text{ mm}^{-1}.
\]

(26.19)

Inverting gives

\[
d_{i}' = \frac{-0.00273 \text{ mm}^{-1}}{1} = -367 \text{ mm}.
\]

(26.20)

The eyepiece’s magnification is thus

\[
m_e = \frac{-d_{i}'}{d_{o}'} = \frac{-367 \text{ mm}}{44.0 \text{ mm}} = 8.33.
\]

(26.21)

So the overall magnification is

\[
m = m_o m_e = (-30.0)(8.33) = -250.
\]

(26.22)

Discussion

Both the objective and the eyepiece contribute to the overall magnification, which is large and negative, consistent with Figure 26.16, where the image is seen to be large and inverted. In this case, the image is virtual and inverted, which cannot happen for a single element (case 2 and case 3 images for single elements are virtual and upright). The final image is 367 mm (0.367 m) to the left of the eyepiece. Had the eyepiece been placed farther from the objective, it could have formed a case 1 image to the right. Such an image could be projected on a screen, but it would be behind the head of the person in the figure and not appropriate for direct viewing. The procedure used to solve this example is applicable in any multiple-element system. Each element is treated in turn, with each forming an image that becomes the object for the next element. The process is not more difficult than for single lenses or mirrors, only lengthier.

Normal optical microscopes can magnify up to 1500\(\times\) with a theoretical resolution of \(0.2 \mu \text{m}\). The lenses can be quite complicated and are composed of multiple elements to reduce aberrations. Microscope objective lenses are particularly important as they primarily gather light from the specimen. Three parameters describe microscope objectives: the numerical aperture (NA), the magnification (\(m\)), and the working distance. The \(NA\) is related to the light gathering ability of a lens and is obtained using the angle of acceptance \(\theta\) formed by the maximum cone of rays focusing on the specimen (see Figure 26.17(a)) and is given by

\[
NA = n \sin \alpha,
\]

(26.23)

where \(n\) is the refractive index of the medium between the lens and the specimen and \(\alpha = \theta/2\). As the angle of acceptance given by \(\theta\) increases, \(NA\) becomes larger and more light is gathered from a smaller focal region giving higher resolution. A 0.75\(NA\) objective gives more detail than a 0.10\(NA\) objective.
Figure 26.17 (a) The numerical aperture \( (NA) \) of a microscope objective lens refers to the light-gathering ability of the lens and is calculated using half the angle of acceptance \( \theta \). (b) Here, \( \alpha \) is half the acceptance angle for light rays from a specimen entering a camera lens, and \( D \) is the diameter of the aperture that controls the light entering the lens.

While the numerical aperture can be used to compare resolutions of various objectives, it does not indicate how far the lens could be from the specimen. This is specified by the “working distance,” which is the distance (in mm usually) from the front lens element of the objective to the specimen, or cover glass. The higher the \( NA \) the closer the lens will be to the specimen and the more chances there are of breaking the cover slip and damaging both the specimen and the lens. The focal length of an objective lens is different than the working distance. This is because objective lenses are made of a combination of lenses and the focal length is measured from inside the barrel. The working distance is a parameter that microscopists can use more readily as it is measured from the outermost lens. The working distance decreases as the \( NA \) and magnification both increase.

The term \( f / \# \) in general is called the \( f \)-number and is used to denote the light per unit area reaching the image plane. In photography, an image of an object at infinity is formed at the focal point and the \( f \)-number is given by the ratio of the focal length \( f \) of the lens and the diameter \( D \) of the aperture controlling the light into the lens (see Figure 26.17(b)). If the acceptance angle is small the \( NA \) of the lens can also be used as given below.

\[
f / \# = \frac{f}{D} \approx \frac{1}{2NA}.
\]  

(26.24)

As the \( f \)-number decreases, the camera is able to gather light from a larger angle, giving wide-angle photography. As usual there is a trade-off. A greater \( f / \# \) means less light reaches the image plane. A setting of \( f / 16 \) usually allows one to take pictures in bright sunlight as the aperture diameter is small. In optical fibers, light needs to be focused into the fiber. Figure 26.18 shows the angle used in calculating the \( NA \) of an optical fiber.

Can the \( NA \) be larger than 1.00? The answer is ‘yes’ if we use immersion lenses in which a medium such as oil, glycerine or water is placed between the objective and the microscope cover slip. This minimizes the mismatch in refractive indices as light rays go through different media, generally providing a greater light-gathering ability and an increase in resolution. Figure 26.19 shows light rays when using air and immersion lenses.
Figure 26.19 Light rays from a specimen entering the objective. Paths for immersion medium of air (a), water (b) \( n = 1.33 \), and oil (c) \( n = 1.51 \) are shown. The water and oil immersions allow more rays to enter the objective, increasing the resolution.

When using a microscope we do not see the entire extent of the sample. Depending on the eyepiece and objective lens we see a restricted region which we say is the field of view. The objective is then manipulated in two-dimensions above the sample to view other regions of the sample. Electronic scanning of either the objective or the sample is used in scanning microscopy. The image formed at each point during the scanning is combined using a computer to generate an image of a larger region of the sample at a selected magnification.

When using a microscope, we rely on gathering light to form an image. Hence most specimens need to be illuminated, particularly at higher magnifications, when observing details that are so small that they reflect only small amounts of light. To make such objects easily visible, the intensity of light falling on them needs to be increased. Special illuminating systems called condensers are used for this purpose. The type of condenser that is suitable for an application depends on how the specimen is examined, whether by transmission, scattering or reflecting. See Figure 26.20 for an example of each. White light sources are common and lasers are often used. Laser light illumination tends to be quite intense and it is important to ensure that the light does not result in the degradation of the specimen.

Figure 26.20 Illumination of a specimen in a microscope. (a) Transmitted light from a condenser lens. (b) Transmitted light from a mirror condenser. (c) Dark field illumination by scattering (the illuminating beam misses the objective lens). (d) High magnification illumination with reflected light – normally laser light.

We normally associate microscopes with visible light but x-ray and electron microscopes provide greater resolution. The focusing and basic physics is the same as that just described, even though the lenses require different technology. The electron microscope requires vacuum chambers so that the electrons can proceed unheeded. Magnifications of 50 million times provide the ability to determine positions of individual atoms within materials. An electron microscope is shown in Figure 26.21. We do not use our eyes to form images; rather images are recorded electronically and displayed on computers. In fact observing and saving images formed by optical microscopes on computers is now done routinely. Video recordings of what occurs
in a microscope can be made for viewing by many people at later dates. Physics provides the science and tools needed to generate the sequence of time-lapse images of meiosis similar to the sequence sketched in Figure 26.22.

![Figure 26.21](image)

**Figure 26.21** An electron microscope has the capability to image individual atoms on a material. The microscope uses vacuum technology, sophisticated detectors and state of the art image processing software. (credit: Dave Pape)

![Figure 26.22](image)

**Figure 26.22** The image shows a sequence of events that takes place during meiosis. (credit: PatríciaR, Wikimedia Commons; National Center for Biotechnology Information)

---

**Take-Home Experiment: Make a Lens**

Look through a clear glass or plastic bottle and describe what you see. Now fill the bottle with water and describe what you see. Use the water bottle as a lens to produce the image of a bright object and estimate the focal length of the water bottle lens. How is the focal length a function of the depth of water in the bottle?

---

**26.5 Telescopes**

Telescopes are meant for viewing distant objects, producing an image that is larger than the image that can be seen with the unaided eye. Telescopes gather far more light than the eye, allowing dim objects to be observed with greater magnification and better resolution. Although Galileo is often credited with inventing the telescope, he actually did not. What he did was more important. He constructed several early telescopes, was the first to study the heavens with them, and made monumental discoveries using them. Among these are the moons of Jupiter, the craters and mountains on the Moon, the details of sunspots, and the fact that the Milky Way is composed of vast numbers of individual stars.

**Figure 26.23(a)** shows a telescope made of two lenses, the convex objective and the concave eyepiece, the same construction used by Galileo. Such an arrangement produces an upright image and is used in spyglasses and opera glasses.
Galileo made telescopes with a convex objective and a concave eyepiece. These produce an upright image and are used in spyglasses. Most simple telescopes have two convex lenses. The objective forms a case 1 image that is the object for the eyepiece. The eyepiece forms a case 2 final image that is magnified. The most common two-lens telescope, like the simple microscope, uses two convex lenses and is shown in Figure 26.23(b). The object is so far away from the telescope that it is essentially at infinity compared with the focal lengths of the lenses \(d_o \approx \infty\). The first image is thus produced at \(d_i = f_o\), as shown in the figure. To prove this, note that

\[
\frac{1}{d_i} = \frac{1}{f_o} - \frac{1}{f_e} = \frac{1}{f_o} - \frac{1}{\infty}.
\]

Because \(1/\infty = 0\), this simplifies to

\[
\frac{1}{d_i} = \frac{1}{f_o},
\]

which implies that \(d_i = f_o\), as claimed. It is true that for any distant object and any lens or mirror, the image is at the focal length.

The first image formed by a telescope objective as seen in Figure 26.23(b) will not be large compared with what you might see by looking at the object directly. For example, the spot formed by sunlight focused on a piece of paper by a magnifying glass is the image of the Sun, and it is small. The telescope eyepiece (like the microscope eyepiece) magnifies this first image. The distance between the eyepiece and the objective lens is made slightly less than the sum of their focal lengths so that the first image is closer to the eyepiece than its focal length. That is, \(d_o'\) is less than \(f_e\), and so the eyepiece forms a case 2 image that is large and to the left for easy viewing. If the angle subtended by an object as viewed by the unaided eye is \(\theta\), and the angle subtended by the telescope image is \(\theta'\), then the angular magnification \(M\) is defined to be their ratio. That is, \(M = \theta' / \theta\). It can be shown that the angular magnification of a telescope is related to the focal lengths of the objective and eyepiece; and is given by

\[
M = \frac{\theta'}{\theta} = -\frac{f_o}{f_e}.
\]

The minus sign indicates the image is inverted. To obtain the greatest angular magnification, it is best to have a long focal length objective and a short focal length eyepiece. The greater the angular magnification \(M\), the larger an object will appear when viewed through a telescope, making more details visible. Limits to observable details are imposed by many factors, including lens quality and atmospheric disturbance.

The image in most telescopes is inverted, which is unimportant for observing the stars but a real problem for other applications, such as telescopes on ships or telescopic gun sights. If an upright image is needed, Galileo’s arrangement in Figure 26.23(a) can be used. But a more common arrangement is to use a third convex lens as an eyepiece, increasing the distance between the first two and inverting the image once again as seen in Figure 26.24.
A telescope can also be made with a concave mirror as its first element or objective, since a concave mirror acts like a convex lens as seen in Figure 26.25. Flat mirrors are often employed in optical instruments to make them more compact or to send light to cameras and other sensing devices. There are many advantages to using mirrors rather than lenses for telescope objectives. Mirrors can be constructed much larger than lenses and can, thus, gather large amounts of light, as needed to view distant galaxies, for example. Large and relatively flat mirrors have very long focal lengths, so that great angular magnification is possible.

Telescopes, like microscopes, can utilize a range of frequencies from the electromagnetic spectrum. Figure 26.26(a) shows the Australia Telescope Compact Array, which uses six 22-m antennas for mapping the southern skies using radio waves. Figure 26.26(b) shows the focusing of x rays on the Chandra X-ray Observatory—a satellite orbiting earth since 1999 and looking at high temperature events as exploding stars, quasars, and black holes. X rays, with much more energy and shorter wavelengths than RF and light, are mainly absorbed and not reflected when incident perpendicular to the medium. But they can be reflected when incident at small glancing angles, much like a rock will skip on a lake if thrown at a small angle. The mirrors for the Chandra consist of a long barrelled pathway and 4 pairs of mirrors to focus the rays at a point 10 meters away from the entrance. The mirrors are extremely smooth and consist of a glass ceramic base with a thin coating of metal (iridium). Four pairs of precision manufactured mirrors are exquisitely shaped and aligned so that x rays ricochet off the mirrors like bullets off a wall, focusing on a spot.
26.6 Aberrations

Real lenses behave somewhat differently from how they are modeled using the thin lens equations, producing aberrations. An aberration is a distortion in an image. There are a variety of aberrations due to a lens size, material, thickness, and position of the object. One common type of aberration is chromatic aberration, which is related to color. Since the index of refraction of lenses depends on color or wavelength, images are produced at different places and with different magnifications for different colors. (The law of reflection is independent of wavelength, and so mirrors do not have this problem. This is another advantage for mirrors in optical systems such as telescopes.) Figure 26.28(a) shows chromatic aberration for a single convex lens and its partial correction with a two-lens system. Violet rays are bent more than red, since they have a higher index of refraction and are thus focused closer to the lens. The diverging lens partially corrects this, although it is usually not possible to do so completely. Lenses of different materials and having different dispersions may be used. For example an achromatic doublet consisting of a converging lens made of crown glass and a diverging lens made of flint glass in contact can dramatically reduce chromatic aberration (see Figure 26.28(b)).

Quite often in an imaging system the object is off-center. Consequently, different parts of a lens or mirror do not refract or reflect the image to the same point. This type of aberration is called a coma and is shown in Figure 26.29. The image in this case often appears pear-shaped. Another common aberration is spherical aberration where rays converging from the outer edges of a lens converge to a focus closer to the lens and rays...
closer to the axis focus further (see Figure 26.30). Aberrations due to astigmatism in the lenses of the eyes are discussed in Vision Correction, and a chart used to detect astigmatism is shown in Figure 26.8. Such aberrations and can also be an issue with manufactured lenses.

Figure 26.28 (a) Chromatic aberration is caused by the dependence of a lens’s index of refraction on color (wavelength). The lens is more powerful for violet (V) than for red (R), producing images with different locations and magnifications. (b) Multiple-lens systems can partially correct chromatic aberrations, but they may require lenses of different materials and add to the expense of optical systems such as cameras.

Figure 26.29 A coma is an aberration caused by an object that is off-center, often resulting in a pear-shaped image. The rays originate from points that are not on the optical axis and they do not converge at one common focal point.

Figure 26.30 Spherical aberration is caused by rays focusing at different distances from the lens.

The image produced by an optical system needs to be bright enough to be discerned. It is often a challenge to obtain a sufficiently bright image. The brightness is determined by the amount of light passing through the optical system. The optical components determining the brightness are the diameter of the lens and the diameter of pupils, diaphragms or aperture stops placed in front of lenses. Optical systems often have entrance and exit pupils to specifically reduce aberrations but they inevitably reduce brightness as well. Consequently, optical systems need to strike a balance between the various components used. The iris in the eye dilates and constricts, acting as an entrance pupil. You can see objects more clearly by looking through a small hole made with your hand in the shape of a fist. Squinting, or using a small hole in a piece of paper, also will make the object sharper.

So how are aberrations corrected? The lenses may also have specially shaped surfaces, as opposed to the simple spherical shape that is relatively easy to produce. Expensive camera lenses are large in diameter, so that they can gather more light, and need several elements to correct for various aberrations. Further, advances in materials science have resulted in lenses with a range of refractive indices—technically referred to as graded index (GRIN) lenses. Spectacles often have the ability to provide a range of focusing ability using similar techniques. GRIN lenses are particularly important at the end of optical fibers in endoscopes. Advanced computing techniques allow for a range of corrections on images after the image has been collected and certain characteristics of the optical system are known. Some of these techniques are sophisticated versions of what are available on commercial packages like Adobe Photoshop.
aberration: failure of rays to converge at one focus because of limitations or defects in a lens or mirror
accommodation: the ability of the eye to adjust its focal length is known as accommodation
adaptive optics: optical technology in which computers adjust the lenses and mirrors in a device to correct for image distortions
angular magnification: a ratio related to the focal lengths of the objective and eyepiece and given as \( M = -\frac{f_o}{f_e} \)
astigmatism: the result of an inability of the cornea to properly focus an image onto the retina
color constancy: a part of the visual perception system that allows people to perceive color in a variety of conditions and to see some consistency in the color
compound microscope: a microscope constructed from two convex lenses, the first serving as the ocular lens (close to the eye) and the second serving as the objective lens
eyepiece: the lens or combination of lenses in an optical instrument nearest to the eye of the observer
far point: the object point imaged by the eye onto the retina in an unaccommodated eye
farsightedness: another term for hyperopia, the condition of an eye where incoming rays of light reach the retina before they converge into a focused image
hues: identity of a color as it relates specifically to the spectrum
hyperopia: the condition of an eye where incoming rays of light reach the retina before they converge into a focused image
laser vision correction: a medical procedure used to correct astigmatism and eyesight deficiencies such as myopia and hyperopia
myopia: a visual defect in which distant objects appear blurred because their images are focused in front of the retina rather than being focused on the retina
near point: the point nearest the eye at which an object is accurately focused on the retina at full accommodation
nearsightedness: another term for myopia, a visual defect in which distant objects appear blurred because their images are focused in front of the retina rather than being focused on the retina
numerical aperture: a number or measure that expresses the ability of a lens to resolve fine detail in an object being observed. Derived by mathematical formula
\[
NA = n \sin \alpha,
\]
where \( n \) is the refractive index of the medium between the lens and the specimen and \( \alpha = \frac{\theta}{2} \)
objective lens: the lens nearest to the object being examined
presbyopia: a condition in which the lens of the eye becomes progressively unable to focus on objects close to the viewer
retinex theory of color vision: the ability to perceive color in an ambient-colored environment
retinex: a theory proposed to explain color and brightness perception and constancies; is a combination of the words retina and cortex, which are the two areas responsible for the processing of visual information
rods and cones: two types of photoreceptors in the human retina; rods are responsible for vision at low light levels, while cones are active at higher light levels
simplified theory of color vision: a theory that states that there are three primary colors, which correspond to the three types of cones

Section Summary

26.1 Physics of the Eye
• Image formation by the eye is adequately described by the thin lens equations:
\[
P = \frac{1}{d_o} + \frac{1}{d_i} \quad \text{and} \quad \frac{h_i}{h_o} = -\frac{d_i}{d_o} = m.
\]
• The eye produces a real image on the retina by adjusting its focal length and power in a process called accommodation.
• For close vision, the eye is fully accommodated and has its greatest power, whereas for distant vision, it is totally relaxed and has its smallest power.
• The loss of the ability to accommodate with age is called presbyopia, which is corrected by the use of a converging lens to add power for close vision.

26.2 Vision Correction
• Nearsightedness, or myopia, is the inability to see distant objects and is corrected with a diverging lens to reduce power.
• Farsightedness, or hyperopia, is the inability to see close objects and is corrected with a converging lens to increase power.
In myopia and hyperopia, the corrective lenses produce images at a distance that the person can see clearly—the far point and near point, respectively.

26.3 Color and Color Vision
- The eye has four types of light receptors—rods and three types of color-sensitive cones.
- The rods are good for night vision, peripheral vision, and motion changes, while the cones are responsible for central vision and color.
- We perceive many hues, from light having mixtures of wavelengths.
- A simplified theory of color vision states that there are three primary colors, which correspond to the three types of cones, and that various combinations of the primary colors produce all the hues.
- The true color of an object is related to its relative absorption of various wavelengths of light. The color of a light source is related to the wavelengths it produces.
- Color constancy is the ability of the eye-brain system to discern the true color of an object illuminated by various light sources.
- The retinex theory of color vision explains color constancy by postulating the existence of three retinexes or image systems, associated with the three types of cones that are compared to obtain sophisticated information.

26.4 Microscopes
- The microscope is a multiple-element system having more than a single lens or mirror.
- Many optical devices contain more than a single lens or mirror. These are analysed by considering each element sequentially. The image formed by the first is the object for the second, and so on. The same ray tracing and thin lens techniques apply to each lens element.
- The overall magnification of a multiple-element system is the product of the magnifications of its individual elements. For a two-element system with an objective and an eyepiece, this is
  \[ m = m_o m_e, \]
  where \( m_o \) is the magnification of the objective and \( m_e \) is the magnification of the eyepiece, such as for a microscope.
- Microscopes are instruments for allowing us to see detail we would not be able to see with the unaided eye and consist of a range of components.
- The eyepiece and objective contribute to the magnification. The numerical aperture (NA) of an objective is given by
  \[ NA = n \sin \alpha \]
  where \( n \) is the refractive index and \( \alpha \) the angle of acceptance.
- Immersion techniques are often used to improve the light gathering ability of microscopes. The specimen is illuminated by transmitted, scattered or reflected light though a condenser.
- The \( f/\# \) describes the light gathering ability of a lens. It is given by
  \[ f/# = \frac{f}{D} \approx \frac{1}{2NA}. \]

26.5 Telescopes
- Simple telescopes can be made with two lenses. They are used for viewing objects at large distances and utilize the entire range of the electromagnetic spectrum.
- The angular magnification \( M \) for a telescope is given by
  \[ M = \frac{\phi'}{\phi} = \frac{f_o}{f_e}, \]
  where \( \phi \) is the angle subtended by an object viewed by the unaided eye, \( \phi' \) is the angle subtended by a magnified image, and \( f_o \) and \( f_e \) are the focal lengths of the objective and the eyepiece.

26.6 Aberrations
- Aberrations or image distortions can arise due to the finite thickness of optical instruments, imperfections in the optical components, and limitations on the ways in which the components are used.
- The means for correcting aberrations range from better components to computational techniques.

Conceptual Questions

26.1 Physics of the Eye
1. If the lens of a person's eye is removed because of cataracts (as has been done since ancient times), why would you expect a spectacle lens of about 16 D to be prescribed?
2. A cataract is cloudiness in the lens of the eye. Is light dispersed or diffused by it?
3. When laser light is shone into a relaxed normal-vision eye to repair a tear by spot-welding the retina to the back of the eye, the rays entering the eye must be parallel. Why?
4. How does the power of a dry contact lens compare with its power when resting on the tear layer of the eye? Explain.
5. Why is your vision so blurry when you open your eyes while swimming under water? How does a face mask enable clear vision?

26.2 Vision Correction
6. It has become common to replace the cataract-clouded lens of the eye with an internal lens. This intraocular lens can be chosen so that the person has perfect distant vision. Will the person be able to read without glasses? If the person was nearsighted, is the power of the intraocular lens greater or less than the removed lens?
7. If the cornea is to be reshaped (this can be done surgically or with contact lenses) to correct myopia, should its curvature be made greater or smaller? Explain. Also explain how hyperopia can be corrected.
8. If there is a fixed percent uncertainty in LASIK reshaping of the cornea, why would you expect those people with the greatest correction to have a poorer chance of normal distant vision after the procedure?

9. A person with presbyopia has lost some or all of the ability to accommodate the power of the eye. If such a person's distant vision is corrected with LASIK, will she still need reading glasses? Explain.

### 26.3 Color and Color Vision

10. A pure red object on a black background seems to disappear when illuminated with pure green light. Explain why.

11. What is color constancy, and what are its limitations?

12. There are different types of color blindness related to the malfunction of different types of cones. Why would it be particularly useful to study those rare individuals who are color blind only in one eye or who have a different type of color blindness in each eye?

13. Propose a way to study the function of the rods alone, given they can sense light about 1000 times dimmer than the cones.

### 26.4 Microscopes

14. Geometric optics describes the interaction of light with macroscopic objects. Why, then, is it correct to use geometric optics to analyse a microscope's image?

15. The image produced by the microscope in Figure 26.16 cannot be projected. Could extra lenses or mirrors project it? Explain.

16. Why not have the objective of a microscope form a case 2 image with a large magnification? (Hint: Consider the location of that image and the difficulty that would pose for using the eyepiece as a magnifier.)

17. What advantages do oil immersion objectives offer?

18. How does the $NA$ of a microscope compare with the $NA$ of an optical fiber?

### 26.5 Telescopes

19. If you want your microscope or telescope to project a real image onto a screen, how would you change the placement of the eyepiece relative to the objective?

### 26.6 Aberrations

20. List the various types of aberrations. What causes them and how can each be reduced?
26.1 Physics of the Eye

Unless otherwise stated, the lens-to-retina distance is 2.00 cm.

1. What is the power of the eye when viewing an object 50.0 cm away?
2. Calculate the power of the eye when viewing an object 3.00 m away.
3. (a) The print in many books averages 3.50 mm in height. How high is the image of the print on the retina when the book is held 30.0 cm from the eye?
   (b) Compare the size of the print to the sizes of rods and cones in the fovea and discuss the possible details observable in the letters. (The eye-brain system can perform better because of interconnections and higher order image processing.)
4. Suppose a certain person’s visual acuity is such that he can see objects clearly that form an image 4.00 μm high on his retina. What is the maximum distance at which he can read the 75.0 cm high letters on the side of an airplane?
5. People who do very detailed work close up, such as jewellers, often can see objects clearly at much closer distance than the normal 25 cm. (a) What is the power of the eyes of a woman who can see an object clearly at a distance of only 8.00 cm?
   (b) What is the size of an image of a 1.00 mm object, such as lettering inside a ring, held at this distance?
   (c) What would the size of the image be if the object were held at the normal 25.0 cm distance?

26.2 Vision Correction

6. What is the far point of a person whose eyes have a relaxed power of 50.5 D?
7. What is the near point of a person whose eyes have an accommodated power of 53.5 D?
8. (a) A laser vision correction reshaping the cornea of a myopic patient reduces the power of his eye by 9.00 D, with a ±5.0% uncertainty in the final correction. What is the range of diopters for spectacle lenses that this person might need after LASIK procedure? (b) Was the person nearsighted or farsighted before the procedure? How do you know?
9. In a LASIK vision correction, the power of a patient’s eye is increased by 3.00 D. Assuming this produces normal close vision, what was the patient’s near point before the procedure?
10. What was the previous far point of a patient who had laser vision correction that reduced the power of her eye by 7.00 D, producing normal distant vision for her?
11. A severely myopic patient has a far point of 5.00 cm. By how many diopters should the power of his eye be reduced in laser vision correction to obtain normal distant vision for him?
12. A student’s eyes, while reading the blackboard, have a power of 51.0 D. How far is the board from his eyes?
13. The power of a physician’s eyes is 53.0 D while examining a patient. How far from her eyes is the feature being examined?
14. A young woman with normal distant vision has a 10.0% ability to accommodate (that is, increase) the power of her eyes. What is the closest object she can see clearly?
15. The far point of a myopic administrator is 50.0 cm. (a) What is the relaxed power of his eyes? (b) If he has the normal 8.00% ability to accommodate, what is the closest object he can see clearly?
16. A very myopic man has a far point of 20.0 cm. What power contact lens (when on the eye) will correct his distant vision?
17. Repeat the previous problem for eyeglasses held 1.50 cm from the eyes.
18. A myopic person sees that her contact lens prescription is −4.00 D. What is her far point?

19. Repeat the previous problem for glasses that are 1.75 cm from the eyes.
20. The contact lens prescription for a mildly farsighted person is 0.750 D, and the person has a near point of 29.0 cm. What is the power of the tear layer between the cornea and the lens if the correction is ideal, taking the tear layer into account?
21. A nearsighted man cannot see objects clearly beyond 20 cm from his eyes. How close must he stand to a mirror in order to see what he is doing when he shaves?
22. A mother sees that her child’s contact lens prescription is 0.750 D. What is the child’s near point?
23. Repeat the previous problem for glasses that are 2.20 cm from the eyes.
24. The contact lens prescription for a nearsighted person is −4.00 D and the person has a far point of 22.5 cm. What is the power of the tear layer between the cornea and the lens if the correction is ideal, taking the tear layer into account?

25. Unreasonable Results

A boy has a near point of 50 cm and a far point of 500 cm. Will a −4.00 D lens correct his far point to infinity?

26.4 Microscopes

26. A microscope with an overall magnification of 800 has an objective that magnifies by 200. (a) What is the magnification of the eyepiece? (b) If there are two other objectives that can be used, having magnifications of 100 and 400, what other total magnifications are possible?
27. (a) What magnification is produced by a 0.150 cm focal length microscope objective that is 0.155 cm from the object being viewed? (b) What is the overall magnification if an 8× eyepiece (one that produces a magnification of 8.00) is used?
28. (a) Where does an object need to be placed relative to a microscope for its 0.500 cm focal length objective to produce a magnification of −400 ? (b) Where should the 5.00 cm focal length eyepiece be placed to produce a further fourfold (4.00) magnification?
29. You switch from a 1.40× eyepiece oil immersion objective to a 1.40× oil immersion objective. What are the acceptance angles for each? Compare and comment on the values. Which would you use first to locate the target area on your specimen?
30. An amoeba is 0.305 cm away from the 0.300 cm focal length objective lens of a microscope. (a) Where is the image formed by the objective lens? (b) What is this image’s magnification? (c) An eyepiece with a 2.00 cm focal length is placed 20.0 cm from the objective. Where is the final image? (d) What magnification is produced by the eyepiece? (e) What is the overall magnification? (See Figure 26.16.)
31. You are using a standard microscope with a 0.10× 4× objective and switch to a 0.65× 40× objective. What are the acceptance angles for each? Compare and comment on the values. Which would you use first to locate the target area on of your specimen? (See Figure 26.17.)

32. Unreasonable Results

Your friends show you an image through a microscope. They tell you that the microscope has an objective with a 0.500 cm focal length and an eyepiece with a 5.00 cm focal length. The resulting overall magnification is 250,000. Are these viable values for a microscope?

26.5 Telescopes

Unless otherwise stated, the lens-to-retina distance is 2.00 cm.

33. What is the angular magnification of a telescope that has a 100 cm focal length objective and a 2.50 cm focal length eyepiece?
34. Find the distance between the objective and eyepiece lenses in the telescope in the above problem needed to produce a final image very far
35. A large reflecting telescope has an objective mirror with a 10.0 m radius of curvature. What angular magnification does it produce when a 3.00 m focal length eyepiece is used?

36. A small telescope has a concave mirror with a 2.00 m radius of curvature for its objective. Its eyepiece is a 4.00 cm focal length lens. (a) What is the telescope’s angular magnification? (b) What angle is subtended by a 25,000 km diameter sunspot? (c) What is the angle of its telescopic image?

37. A 7.5x binocular produces an angular magnification of $-7.50$. acting like a telescope. (Mirrors are used to make the image upright.) If the binoculars have objective lenses with a 75.0 cm focal length, what is the focal length of the eyepiece lenses?

38. Construct Your Own Problem
Consider a telescope of the type used by Galileo, having a convex objective and a concave eyepiece as illustrated in Figure 26.23(a). Construct a problem in which you calculate the location and size of the image produced. Among the things to be considered are the focal lengths of the lenses and their relative placements as well as the size and location of the object. Verify that the angular magnification is greater than one. That is, the angle subtended at the eye by the image is greater than the angle subtended by the object.

26.6 Aberrations

39. Integrated Concepts
(a) During laser vision correction, a brief burst of 193 nm ultraviolet light is projected onto the cornea of the patient. It makes a spot 1.00 mm in diameter and deposits 0.500 mJ of energy. Calculate the depth of the layer ablated, assuming the corneal tissue has the same properties as water and is initially at 34.0°C. The tissue’s temperature is increased to 100°C and evaporated without further temperature increase.

(b) Does your answer imply that the shape of the cornea can be finely controlled?
Learning Objectives

27.1. The Wave Aspect of Light: Interference
• Discuss the wave character of light.
• Identify the changes when light enters a medium.

27.2. Huygens’s Principle: Diffraction
• Discuss the propagation of transverse waves.
• Discuss Huygens’s principle.
• Explain the bending of light.

27.3. Young’s Double Slit Experiment
• Explain the phenomena of interference.
• Define constructive interference for a double slit and destructive interference for a double slit.

27.4. Multiple Slit Diffraction
• Discuss the pattern obtained from diffraction grating.
• Explain diffraction grating effects.

27.5. Single Slit Diffraction
• Discuss the single slit diffraction pattern.

27.6. Limits of Resolution: The Rayleigh Criterion
• Discuss the Rayleigh criterion.

27.7. Thin Film Interference
• Discuss the rainbow formation by thin films.

27.8. Polarization
• Discuss the meaning of polarization.
• Discuss the property of optical activity of certain materials.

27.9. *Extended Topic* Microscopy Enhanced by the Wave Characteristics of Light
• Discuss the different types of microscopes.
Introduction to Wave Optics

Examine a compact disc under white light, noting the colors observed and locations of the colors. Determine if the spectra are formed by diffraction from circular lines centered at the middle of the disc and, if so, what is their spacing. If not, determine the type of spacing. Also with the CD, explore the spectra of a few light sources, such as a candle flame, incandescent bulb, halogen light, and fluorescent light. Knowing the spacing of the rows of pits in the compact disc, estimate the maximum spacing that will allow the given number of megabytes of information to be stored.

If you have ever looked at the reds, blues, and greens in a sunlit soap bubble and wondered how straw-colored soapy water could produce them, you have hit upon one of the many phenomena that can only be explained by the wave character of light (see Figure 27.2). The same is true for the colors seen in an oil slick or in the light reflected from a compact disc. These and other interesting phenomena, such as the dispersion of white light into a rainbow of colors when passed through a narrow slit, cannot be explained fully by geometric optics. In these cases, light interacts with small objects and exhibits its wave characteristics. The branch of optics that considers the behavior of light when it exhibits wave characteristics (particularly when it interacts with small objects) is called wave optics (sometimes called physical optics). It is the topic of this chapter.

Figure 27.2 These soap bubbles exhibit brilliant colors when exposed to sunlight. How are the colors produced if they are not pigments in the soap? (credit: Scott Robinson, Flickr)

27.1 The Wave Aspect of Light: Interference

We know that visible light is the type of electromagnetic wave to which our eyes respond. Like all other electromagnetic waves, it obeys the equation

\[ c = f \lambda, \]

where \( c = 3 \times 10^8 \text{ m/s} \) is the speed of light in vacuum, \( f \) is the frequency of the electromagnetic waves, and \( \lambda \) is its wavelength. The range of visible wavelengths is approximately 380 to 760 nm. As is true for all waves, light travels in straight lines and acts like a ray when it interacts with objects several times as large as its wavelength. However, when it interacts with smaller objects, it displays its wave characteristics prominently. Interference is the hallmark of a wave, and in Figure 27.3 both the ray and wave characteristics of light can be seen. The laser beam emitted by the observatory epitomizes a ray, traveling in a straight line. However, passing a pure-wavelength beam through vertical slits with a size close to the wavelength of the beam reveals the wave character of light, as the beam spreads out horizontally into a pattern of bright and dark regions caused by systematic constructive and destructive interference. Rather than spreading out, a ray would continue traveling straight ahead after passing through slits.

Making Connections: Waves

The most certain indication of a wave is interference. This wave characteristic is most prominent when the wave interacts with an object that is not large compared with the wavelength. Interference is observed for water waves, sound waves, light waves, and (as we will see in Special Relativity) for matter waves, such as electrons scattered from a crystal.
Light has wave characteristics in various media as well as in a vacuum. When light goes from a vacuum to some medium, like water, its speed and wavelength change, but its frequency $f$ remains the same. (We can think of light as a forced oscillation that must have the frequency of the original source.) The speed of light in a medium is $v = c/n$, where $n$ is its index of refraction. If we divide both sides of equation $c = f\lambda$ by $n$, we get $c/n = v = f\lambda/n$. This implies that $v = f\lambda_n$, where $\lambda_n$ is the wavelength in a medium and that

$$\lambda_n = \frac{\lambda}{n},$$  

where $\lambda$ is the wavelength in vacuum and $n$ is the medium’s index of refraction. Therefore, the wavelength of light is smaller in any medium than it is in vacuum. In water, for example, which has $n = 1.333$, the range of visible wavelengths is $(380 \text{ nm})/1.333$ to $(760 \text{ nm})/1.333$, or $\lambda_n = 285$ to $570$ nm. Although wavelengths change while traveling from one medium to another, colors do not, since colors are associated with frequency.

### 27.2 Huygens’s Principle: Diffraction

Figure 27.4 shows how a transverse wave looks as viewed from above and from the side. A light wave can be imagined to propagate like this, although we do not actually see it wiggling through space. From above, we view the wavefronts (or wave crests) as we would by looking down on the ocean waves. The side view would be a graph of the electric or magnetic field. The view from above is perhaps the most useful in developing concepts about wave optics.

![View from above and side](image)

**Figure 27.4** A transverse wave, such as an electromagnetic wave like light, as viewed from above and from the side. The direction of propagation is perpendicular to the wavefronts (or wave crests) and is represented by an arrow like a ray.

The Dutch scientist Christiaan Huygens (1629–1695) developed a useful technique for determining in detail how and where waves propagate. Starting from some known position, Huygens’s principle states that:
Every point on a wavefront is a source of wavelets that spread out in the forward direction at the same speed as the wave itself. The new wavefront is a line tangent to all of the wavelets.

Figure 27.5 shows how Huygens’s principle is applied. A wavefront is the long edge that moves, for example, the crest or the trough. Each point on the wavefront emits a semicircular wave that moves at the propagation speed \( v \). These are drawn at a time \( t \) later, so that they have moved a distance \( s = vt \). The new wavefront is a line tangent to the wavelets and is where we would expect the wave to be a time \( t \) later. Huygens’s principle works for all types of waves, including water waves, sound waves, and light waves. We will find it useful not only in describing how light waves propagate, but also in explaining the laws of reflection and refraction. In addition, we will see that Huygens’s principle tells us how and where light rays interfere.

Figure 27.5 Huygens’s principle applied to a straight wavefront. Each point on the wavefront emits a semicircular wavelet that moves a distance \( s = vt \). The new wavefront is a line tangent to the wavelets.

Figure 27.6 shows how a mirror reflects an incoming wave at an angle equal to the incident angle, verifying the law of reflection. As the wavefront strikes the mirror, wavelets are first emitted from the left part of the mirror and then the right. The wavelets closer to the left have had time to travel farther, producing a wavefront traveling in the direction shown.

Figure 27.6 Huygens’s principle applied to a straight wavefront striking a mirror. The wavelets shown were emitted as each point on the wavefront struck the mirror. The tangent to these wavelets shows that the new wavefront has been reflected at an angle equal to the incident angle. The direction of propagation is perpendicular to the wavefront, as shown by the downward-pointing arrows.

The law of refraction can be explained by applying Huygens’s principle to a wavefront passing from one medium to another (see Figure 27.7). Each wavelet in the figure was emitted when the wavefront crossed the interface between the media. Since the speed of light is smaller in the second medium, the waves do not travel as far in a given time, and the new wavefront changes direction as shown. This explains why a ray changes direction to become closer to the perpendicular when light slows down. Snell’s law can be derived from the geometry in Figure 27.7, but this is left as an exercise for ambitious readers.
What happens when a wave passes through an opening, such as light shining through an open door into a dark room? For light, we expect to see a sharp shadow of the doorway on the floor of the room, and we expect no light to bend around corners into other parts of the room. When sound passes through a door, we expect to hear it everywhere in the room and, thus, expect that sound spreads out when passing through such an opening (see Figure 27.8). What is the difference between the behavior of sound waves and light waves in this case? The answer is that light has very short wavelengths and acts like a ray. Sound has wavelengths on the order of the size of the door and bends around corners (for frequency of 1000 Hz, $\lambda = c/f = (330 \text{ m/s})/(1000 \text{ s}^{-1}) = 0.33 \text{ m}$, about three times smaller than the width of the doorway).

If we pass light through smaller openings, often called slits, we can use Huygens’s principle to see that light bends as sound does (see Figure 27.9). The bending of a wave around the edges of an opening or an obstacle is called diffraction. Diffraction is a wave characteristic and occurs for all types of waves. If diffraction is observed for some phenomenon, it is evidence that the phenomenon is a wave. Thus the horizontal diffraction of the laser beam after it passes through slits in Figure 27.3 is evidence that light is a wave.

27.3 Young’s Double Slit Experiment

Although Christiaan Huygens thought that light was a wave, Isaac Newton did not. Newton felt that there were other explanations for color, and for the interference and diffraction effects that were observable at the time. Owing to Newton’s tremendous stature, his view generally prevailed. The fact that Huygens’s principle worked was not considered evidence that was direct enough to prove that light is a wave. The acceptance of the wave character of light came many years later when, in 1801, the English physicist and physician Thomas Young (1773–1829) did his now-classic double slit experiment (see Figure 27.10).
Why do we not ordinarily observe wave behavior for light, such as observed in Young’s double slit experiment? First, light must interact with something small, such as the closely spaced slits used by Young, to show pronounced wave effects. Furthermore, Young first passed light from a single source (the Sun) through a single slit to make the light somewhat coherent. By coherent, we mean waves are in phase or have a definite phase relationship. Incoherent means the waves have random phase relationships. Why did Young then pass the light through a double slit? The answer to this question is that two slits provide two coherent light sources that then interfere constructively or destructively. Young used sunlight, where each wavelength forms its own pattern, making the effect more difficult to see. We illustrate the double slit experiment with monochromatic (single $\lambda$) light to clarify the effect. Figure 27.11 shows the pure constructive and destructive interference of two waves having the same wavelength and amplitude.

When light passes through narrow slits, it is diffracted into semicircular waves, as shown in Figure 27.12(a). Pure constructive interference occurs where the waves are crest to crest or trough to trough. Pure destructive interference occurs where they are crest to trough. The light must fall on a screen and be scattered into our eyes for us to see the pattern. An analogous pattern for water waves is shown in Figure 27.12(b). Note that regions of constructive and destructive interference move out from the slits at well-defined angles to the original beam. These angles depend on wavelength and the distance between the slits, as we shall see below.
Figure 27.12 Double slits produce two coherent sources of waves that interfere. (a) Light spreads out (diffracts) from each slit, because the slits are narrow. These waves overlap and interfere constructively (bright lines) and destructively (dark regions). We can only see this if the light falls onto a screen and is scattered into our eyes. (b) Double slit interference pattern for water waves are nearly identical to that for light. Wave action is greatest in regions of constructive interference and least in regions of destructive interference. (c) When light that has passed through double slits falls on a screen, we see a pattern such as this. (credit: PASCO)

To understand the double slit interference pattern, we consider how two waves travel from the slits to the screen, as illustrated in Figure 27.13. Each slit is a different distance from a given point on the screen. Thus different numbers of wavelengths fit into each path. Waves start out from the slits in phase (crest to crest), but they may end up out of phase (crest to trough) at the screen if the paths differ in length by half a wavelength, interfering destructively as shown in Figure 27.13(a). If the paths differ by a whole wavelength, then the waves arrive in phase (crest to crest) at the screen, interfering constructively as shown in Figure 27.13(b). More generally, if the paths taken by the two waves differ by any half-integral number of wavelengths $\left(\frac{1}{2}\lambda, \frac{3}{2}\lambda, \frac{5}{2}\lambda, \text{ etc.}\right)$, then destructive interference occurs. Similarly, if the paths taken by the two waves differ by any integral number of wavelengths ($\lambda, 2\lambda, 3\lambda, \text{ etc.}$), then constructive interference occurs.

Take-Home Experiment: Using Fingers as Slits

Look at a light, such as a street lamp or incandescent bulb, through the narrow gap between two fingers held close together. What type of pattern do you see? How does it change when you allow the fingers to move a little farther apart? Is it more distinct for a monochromatic source, such as the yellow light from a sodium vapor lamp, than for an incandescent bulb?

Figure 27.13 Waves follow different paths from the slits to a common point on a screen. (a) Destructive interference occurs here, because one path is a half wavelength longer than the other. The waves start in phase but arrive out of phase. (b) Constructive interference occurs here because one path is a whole wavelength longer than the other. The waves start out and arrive in phase.

Figure 27.14 shows how to determine the path length difference for waves traveling from two slits to a common point on a screen. If the screen is a large distance away compared with the distance between the slits, then the angle $\theta$ between the path and a line from the slits to the screen (see the figure) is nearly the same for each path. The difference between the paths is shown in the figure; simple trigonometry shows it to be $d \sin \theta$, where $d$ is the distance between the slits. To obtain constructive interference for a double slit, the path length difference must be an integral multiple of the wavelength, or

$$d \sin \theta = m\lambda, \text{ for } m = 0, 1, -1, 2, -2, \ldots \quad (\text{constructive}). \quad (27.3)$$

Similarly, to obtain destructive interference for a double slit, the path length difference must be a half-integral multiple of the wavelength, or

$$d \sin \theta = \left(m + \frac{1}{2}\right)\lambda, \text{ for } m = 0, 1, -1, 2, -2, \ldots \quad (\text{destructive}). \quad (27.4)$$

where $\lambda$ is the wavelength of the light, $d$ is the distance between slits, and $\theta$ is the angle from the original direction of the beam as discussed above. We call $m$ the order of the interference. For example, $m = 4$ is fourth-order interference.
The paths from each slit to a common point on the screen differ by an amount \( d \sin \theta \), assuming the distance to the screen is much greater than the distance between slits (not to scale here).

The equations for double slit interference imply that a series of bright and dark lines are formed. For vertical slits, the light spreads out horizontally on either side of the incident beam into a pattern called interference fringes, illustrated in Figure 27.15. The intensity of the bright fringes falls off on either side, being brightest at the center. The closer the slits are, the more is the spreading of the bright fringes. We can see this by examining the equation

\[
d \sin \theta = m\lambda, \quad \text{for } m = 0, 1, -1, 2, -2, \ldots .
\]  

(27.5)

For fixed \( \lambda \) and \( m \), the smaller \( d \) is, the larger \( \theta \) must be, since \( \sin \theta = m\lambda / d \). This is consistent with our contention that wave effects are most noticeable when the object the wave encounters (here, slits a distance \( d \) apart) is small. Small \( d \) gives large \( \theta \), hence a large effect.

Figure 27.15 The interference pattern for a double slit has an intensity that falls off with angle. The photograph shows multiple bright and dark lines, or fringes, formed by light passing through a double slit.

Example 27.1 Finding a Wavelength from an Interference Pattern

Suppose you pass light from a He-Ne laser through two slits separated by 0.0100 mm and find that the third bright line on a screen is formed at an angle of 10.95° relative to the incident beam. What is the wavelength of the light?

Strategy

The third bright line is due to third-order constructive interference, which means that \( m = 3 \). We are given \( d = 0.0100 \) mm and \( \theta = 10.95^\circ \). The wavelength can thus be found using the equation \( d \sin \theta = m\lambda \) for constructive interference.

Solution

The equation is \( d \sin \theta = m\lambda \). Solving for the wavelength \( \lambda \) gives

\[
\lambda = \frac{d \sin \theta}{m}.
\]

(27.6)

Substituting known values yields

\[
\lambda = \frac{(0.0100 \text{ mm})(\sin 10.95^\circ)}{3} = 6.33 \times 10^{-4} \text{ mm} = 633 \text{ nm}.
\]

(27.7)

Discussion

To three digits, this is the wavelength of light emitted by the common He-Ne laser. Not by coincidence, this red color is similar to that emitted by neon lights. More important, however, is the fact that interference patterns can be used to measure wavelength. Young did this for visible
wavelengths. This analytical technique is still widely used to measure electromagnetic spectra. For a given order, the angle for constructive interference increases with $\lambda$, so that spectra (measurements of intensity versus wavelength) can be obtained.

**Example 27.2 Calculating Highest Order Possible**

Interference patterns do not have an infinite number of lines, since there is a limit to how big $m$ can be. What is the highest-order constructive interference possible with the system described in the preceding example?

**Strategy and Concept**

The equation $d \sin \theta = m\lambda$ (for $m = 0, 1, -1, 2, -2, \ldots$) describes constructive interference. For fixed values of $d$ and $\lambda$, the larger $m$ is, the larger $\sin \theta$ is. However, the maximum value that $\sin \theta$ can have is 1, for an angle of $90^\circ$. (Larger angles imply that light goes backward and does not reach the screen at all.) Let us find which $m$ corresponds to this maximum diffraction angle.

**Solution**

Solving the equation $d \sin \theta = m\lambda$ for $m$ gives

$$m = \frac{d \sin \theta}{\lambda}.$$  \hspace{1cm} (27.8)

Taking $\sin \theta = 1$ and substituting the values of $d$ and $\lambda$ from the preceding example gives

$$m = \frac{(0.0100 \text{ mm})(1)}{633 \text{ nm}} \approx 15.8.$$  \hspace{1cm} (27.9)

Therefore, the largest integer $m$ can be is 15, or

$$m = 15.$$  \hspace{1cm} (27.10)

**Discussion**

The number of fringes depends on the wavelength and slit separation. The number of fringes will be very large for large slit separations. However, if the slit separation becomes much greater than the wavelength, the intensity of the interference pattern changes so that the screen has two bright lines cast by the slits, as expected when light behaves like a ray. We also note that the fringes get fainter further away from the center. Consequently, not all 15 fringes may be observable.

**27.4 Multiple Slit Diffraction**

An interesting thing happens if you pass light through a large number of evenly spaced parallel slits, called a **diffraction grating**. An interference pattern is created that is very similar to the one formed by a double slit (see Figure 27.16). A diffraction grating can be manufactured by scratching glass with a sharp tool in a number of precisely positioned parallel lines, with the untouched regions acting like slits. These can be photographically mass produced rather cheaply. Diffraction gratings work both for transmission of light, as in Figure 27.16, and for reflection of light, as on butterfly wings and the Australian opal in Figure 27.17 or the CD pictured in the opening photograph of this chapter, Figure 27.1. In addition to their use as novelty items, diffraction gratings are commonly used for spectroscopic dispersion and analysis of light. What makes them particularly useful is the fact that they form a sharper pattern than double slits do. That is, their bright regions are narrower and brighter, while their dark regions are darker. Figure 27.18 shows idealized graphs demonstrating the sharper pattern. Natural diffraction gratings occur in the feathers of certain birds. Tiny, finger-like structures in regular patterns act as reflection gratings, producing constructive interference that gives the feathers colors not solely due to their pigmentation. This is called iridescence.

![Figure 27.16](image)

*Figure 27.16* A diffraction grating is a large number of evenly spaced parallel slits. (a) Light passing through is diffracted in a pattern similar to a double slit, with bright regions at various angles. (b) The pattern obtained for white light incident on a grating. The central maximum is white, and the higher-order maxima disperse white light into a rainbow of colors.
Figure 27.17 (a) This Australian opal and (b) the butterfly wings have rows of reflectors that act like reflection gratings, reflecting different colors at different angles. (credits: (a) Opals-On-Black.com, via Flickr (b) whologwhy, Flickr)

Figure 27.18 Idealized graphs of the intensity of light passing through a double slit (a) and a diffraction grating (b) for monochromatic light. Maxima can be produced at the same angles, but those for the diffraction grating are narrower and hence sharper. The maxima become narrower and the regions between darker as the number of slits is increased.

The analysis of a diffraction grating is very similar to that for a double slit (see Figure 27.19). As we know from our discussion of double slits in Young’s Double Slit Experiment, light is diffracted by each slit and spreads out after passing through. Rays traveling in the same direction (at an angle \( \theta \) relative to the incident direction) are shown in the figure. Each of these rays travels a different distance to a common point on a screen far away. The rays start in phase, and they can be in or out of phase when they reach a screen, depending on the difference in the path lengths traveled. As seen in the figure, each ray travels a distance \( d \sin \theta \) different from that of its neighbor, where \( d \) is the distance between slits. If this distance equals an integral number of wavelengths, the rays all arrive in phase, and constructive interference (a maximum) is obtained. Thus, the condition necessary to obtain constructive interference for a diffraction grating is

\[
d \sin \theta = m\lambda, \quad \text{for } m = 0, 1, -1, 2, -2, \ldots \text{ (constructive),}
\]

where \( d \) is the distance between slits in the grating, \( \lambda \) is the wavelength of light, and \( m \) is the order of the maximum. Note that this is exactly the same equation as for double slits separated by \( d \). However, the slits are usually closer in diffraction gratings than in double slits, producing fewer maxima at larger angles.
Where are diffraction gratings used? Diffraction gratings are key components of monochromators used, for example, in optical imaging of particular wavelengths from biological or medical samples. A diffraction grating can be chosen to specifically analyze a wavelength emitted by molecules in diseased cells in a biopsy sample or to help excite strategic molecules in the sample with a selected frequency of light. Another vital use is in optical fiber technologies where fibers are designed to provide optimum performance at specific wavelengths. A range of diffraction gratings are available for selecting specific wavelengths for such use.

**Take-Home Experiment: Rainbows on a CD**

The spacing $d$ of the grooves in a CD or DVD can be well determined by using a laser and the equation

$$d \sin \theta = m\lambda,$$

for $m = 0, 1, -1, 2, -2, \ldots$ However, we can still make a good estimate of this spacing by using white light and the rainbow of colors that comes from the interference. Reflect sunlight from a CD onto a wall and use your best judgment of the location of a strongly diffracted color to find the separation $d$.

**Example 27.3 Calculating Typical Diffraction Grating Effects**

Diffraction gratings with 10,000 lines per centimeter are readily available. Suppose you have one, and you send a beam of white light through it to a screen 2.00 m away. (a) Find the angles for the first-order diffraction of the shortest and longest wavelengths of visible light (380 and 760 nm). (b) What is the distance between the ends of the rainbow of visible light produced on the screen for first-order interference? (See Figure 27.20.)
The diffraction grating considered in this example produces a rainbow of colors on a screen a distance $x = 2.00 \text{ m}$ from the grating. The distances along the screen are measured perpendicular to the $x$-direction. In other words, the rainbow pattern extends out of the page.

**Strategy**

The angles can be found using the equation

$$d \sin \theta = m\lambda \quad \text{(for } m = 0, 1, -1, 2, -2, \ldots) \tag{27.12}$$

once a value for the slit spacing $d$ has been determined. Since there are 10,000 lines per centimeter, each line is separated by $1/10,000$ of a centimeter. Once the angles are found, the distances along the screen can be found using simple trigonometry.

**Solution for (a)**

The distance between slits is $d = (1 \text{ cm})/10,000 = 1.00 \times 10^{-4} \text{ cm}$ or $1.00 \times 10^{-6} \text{ m}$. Let us call the two angles $\theta_V$ for violet (380 nm) and $\theta_R$ for red (760 nm). Solving the equation $d \sin \theta = m\lambda$ for $\sin \theta$,

$$\sin \theta = \frac{m\lambda}{d}, \quad \text{where } m = 1 \text{ for first order and } \lambda_V = 380 \text{ nm} = 3.80 \times 10^{-7} \text{ m}.$$  

Substituting these values gives

$$\sin \theta_V = \frac{3.80 \times 10^{-7} \text{ m}}{1.00 \times 10^{-6} \text{ m}} = 0.380. \tag{27.14}$$

Thus the angle $\theta_V$ is

$$\theta_V = \sin^{-1} 0.380 = 22.33^\circ. \tag{27.15}$$

Similarly,

$$\sin \theta_R = \frac{7.60 \times 10^{-7} \text{ m}}{1.00 \times 10^{-6} \text{ m}} \tag{27.16}$$

Thus the angle $\theta_R$ is

$$\theta_R = \sin^{-1} 0.760 = 49.46^\circ. \tag{27.17}$$

Notice that in both equations, we reported the results of these intermediate calculations to four significant figures to use with the calculation in part (b).

**Solution for (b)**

The distances on the screen are labeled $y_V$ and $y_R$ in Figure 27.20. Noting that $\tan \theta = y/x$, we can solve for $y_V$ and $y_R$. That is,

$$y_V = x \tan \theta_V = (2.00 \text{ m})(\tan 22.33^\circ) = 0.815 \text{ m} \tag{27.18}$$

$$y_R = x \tan \theta_R = (2.00 \text{ m})(\tan 49.46^\circ) \approx 1.60 \text{ m}.$$
and

\[ y_R = x \tan \theta_R = (2.00 \text{ m})(\tan 49.46^\circ) = 2.338 \text{ m}. \]  

(27.19)

The distance between them is therefore

\[ y_R - y_V = 1.52 \text{ m}. \]  

(27.20)

**Discussion**

The large distance between the red and violet ends of the rainbow produced from the white light indicates the potential this diffraction grating has as a spectroscopic tool. The more it can spread out the wavelengths (greater dispersion), the more detail can be seen in a spectrum. This depends on the quality of the diffraction grating—it must be very precisely made in addition to having closely spaced lines.

### 27.5 Single Slit Diffraction

Light passing through a single slit forms a diffraction pattern somewhat different from those formed by double slits or diffraction gratings. Figure 27.21 shows a single slit diffraction pattern. Note that the central maximum is larger than those on either side, and that the intensity decreases rapidly on either side. In contrast, a diffraction grating produces evenly spaced lines that dim slowly on either side of center.

![Figure 27.21](image)

Figure 27.21 (a) Single slit diffraction pattern. Monochromatic light passing through a single slit has a central maximum and many smaller and dimmer maxima on either side. The central maximum is six times higher than shown. (b) The drawing shows the bright central maximum and dimmer and thinner maxima on either side.

The analysis of single slit diffraction is illustrated in Figure 27.22. Here we consider light coming from different parts of the same slit. According to Huygens’s principle, every part of the wavefront in the slit emits wavelets. These are like rays that start out in phase and head in all directions. (Each ray is perpendicular to the wavefront of a wavelet.) Assuming the screen is very far away compared with the size of the slit, rays heading toward a common destination are nearly parallel. When they travel straight ahead, as in Figure 27.22(a), they remain in phase, and a central maximum is obtained. However, when rays travel at an angle \( \theta \) relative to the original direction of the beam, each travels a different distance to a common location, and they can arrive in or out of phase. In Figure 27.22(b), the ray from the bottom travels a distance of one wavelength \( \lambda \) farther than the ray from the top. Thus a ray from the center travels a distance \( \lambda / 2 \) farther than the one on the left, arrives out of phase, and interferes destructively. A ray from slightly above the center and one from slightly above the bottom will also cancel one another. In fact, each ray from the slit will have another to interfere destructively, and a minimum in intensity will occur at this angle. There will be another minimum at the same angle to the right of the incident direction of the light.
Figure 27.22 Light passing through a single slit is diffracted in all directions and may interfere constructively or destructively, depending on the angle. The difference in path length for rays from either side of the slit is seen to be $D \sin \theta$.

At the larger angle shown in Figure 27.22(c), the path lengths differ by $\frac{3\lambda}{2}$ for rays from the top and bottom of the slit. One ray travels a distance different from the ray from the bottom and arrives in phase, interfering constructively. Two rays, each from slightly above those two, will also add constructively. Most rays from the slit will have another to interfere with constructively, and a maximum in intensity will occur at this angle. However, all rays do not interfere constructively for this situation, and so the maximum is not as intense as the central maximum. Finally, in Figure 27.22(d), the angle shown is large enough to produce a second minimum. As seen in the figure, the difference in path length for rays from either side of the slit is $D \sin \theta$, and we see that a destructive minimum is obtained when this distance is an integral multiple of the wavelength.

Figure 27.23 A graph of single slit diffraction intensity showing the central maximum to be wider and much more intense than those to the sides. In fact the central maximum is six times higher than shown here.
Thus, to obtain **destructive interference for a single slit**,

\[ D \sin \theta = m\lambda, \text{ for } m = 1, -1, 2, -2, 3, \ldots \text{ (destructive)}, \]  

where \( D \) is the slit width, \( \lambda \) is the light's wavelength, \( \theta \) is the angle relative to the original direction of the light, and \( m \) is the order of the minimum. **Figure 27.23** shows a graph of intensity for single slit interference, and it is apparent that the maxima on either side of the central maximum are much less intense and not as wide. This is consistent with the illustration in **Figure 27.21(b)**.

### Example 27.4 Calculating Single Slit Diffraction

Visible light of wavelength 550 nm falls on a single slit and produces its second diffraction minimum at an angle of \( 45.0^\circ \) relative to the incident direction of the light. (a) What is the width of the slit? (b) At what angle is the first minimum produced?

**Strategy**

From the given information, and assuming the screen is far away from the slit, we can use the equation \( D \sin \theta = m\lambda \) first to find \( D \), and again to find the angle for the first minimum \( \theta_1 \).

**Solution for (a)**

We are given that \( \lambda = 550 \text{ nm} \), \( m = 2 \), and \( \theta_2 = 45.0^\circ \). Solving the equation \( D \sin \theta = m\lambda \) for \( D \) and substituting known values gives

\[ D = \frac{m\lambda}{\sin \theta_2} = \frac{2(550 \text{ nm})}{\sin 45.0^\circ} \]

\[ = 1.56 \times 10^{-6} \text{ m}. \]  

**Solution for (b)**

Solving the equation \( D \sin \theta = m\lambda \) for \( \sin \theta_1 \) and substituting the known values gives

\[ \sin \theta_1 = \frac{m\lambda}{D} = \frac{1(550 \times 10^{-9} \text{ m})}{1.56 \times 10^{-6} \text{ m}}. \]

Thus the angle \( \theta_1 \) is

\[ \theta_1 = \sin^{-1} 0.354 = 20.7^\circ. \]

**Discussion**

We see that the slit is narrow (it is only a few times greater than the wavelength of light). This is consistent with the fact that light must interact with an object comparable in size to its wavelength in order to exhibit significant wave effects such as this single slit diffraction pattern. We also see that the central maximum extends \( 20.7^\circ \) on either side of the original beam, for a width of about \( 41^\circ \). The angle between the first and second minima is only about \( 24^\circ \) \((45.0^\circ - 20.7^\circ)\). Thus the second maximum is only about half as wide as the central maximum.
**27.6 Limits of Resolution: The Rayleigh Criterion**

Light diffracts as it moves through space, bending around obstacles, interfering constructively and destructively. While this can be used as a spectroscopic tool—a diffraction grating disperses light according to wavelength, for example, and is used to produce spectra—diffraction also limits the detail we can obtain in images. Figure 27.25(a) shows the effect of passing light through a small circular aperture. Instead of a bright spot with sharp edges, a spot with a fuzzy edge surrounded by circles of light is obtained. This pattern is caused by diffraction similar to that produced by a single slit. Light from different parts of the circular aperture interferes constructively and destructively. The effect is most noticeable when the aperture is small, but the effect is there for large apertures, too.

![Figure 27.25](image)

(a) Monochromatic light passed through a small circular aperture produces this diffraction pattern. (b) Two point light sources that are close to one another produce overlapping images because of diffraction. (c) If they are closer together, they cannot be resolved or distinguished.

How does diffraction affect the detail that can be observed when light passes through an aperture? Figure 27.25(b) shows the diffraction pattern produced by two point light sources that are close to one another. The pattern is similar to that for a single point source, and it is just barely possible to tell that there are two light sources rather than one. If they were closer together, as in Figure 27.25(c), we could not distinguish them, thus limiting the detail or resolution we can obtain. This limit is an inescapable consequence of the wave nature of light.

There are many situations in which diffraction limits the resolution. The acuity of our vision is limited because light passes through the pupil, the circular aperture of our eye. Be aware that the diffraction-like spreading of light is due to the limited diameter of a light beam, not the interaction with an aperture. Thus light passing through a lens with a diameter $D$ shows this effect and spreads, blurring the image, just as light passing through an aperture of diameter $D$ does. So diffraction limits the resolution of any system having a lens or mirror. Telescopes are also limited by diffraction, because of the finite diameter $D$ of their primary mirror.

**Take-Home Experiment: Resolution of the Eye**

Draw two lines on a white sheet of paper (several mm apart). How far away can you be and still distinguish the two lines? What does this tell you about the size of the eye’s pupil? Can you be quantitative? (The size of an adult’s pupil is discussed in *Physics of the Eye*.)

Just what is the limit? To answer that question, consider the diffraction pattern for a circular aperture, which has a central maximum that is wider and brighter than the maxima surrounding it (similar to a slit) [see Figure 27.26(a)]. It can be shown that, for a circular aperture of diameter $D$, the first minimum in the diffraction pattern occurs at $\theta = \frac{1.22 \lambda}{D}$ (providing the aperture is large compared with the wavelength of light, which is the case for most optical instruments). The accepted criterion for determining the diffraction limit to resolution based on this angle was developed by Lord Rayleigh in the 19th century. The Rayleigh criterion for the diffraction limit to resolution states that two images are just resolvable when the center of the diffraction pattern of one is directly over the first minimum of the diffraction pattern of the other. See Figure 27.26(b). The first minimum is at an angle of $\theta = \frac{1.22 \lambda}{D}$, so that two point objects are just resolvable if they are separated by the angle

$$\theta = \frac{1.22 \lambda}{D}$$  \hspace{1cm} (27.25)

where $\lambda$ is the wavelength of light (or other electromagnetic radiation) and $D$ is the diameter of the aperture, lens, mirror, etc., with which the two objects are observed. In this expression, $\theta$ has units of radians.
Figure 27.26 (a) Graph of intensity of the diffraction pattern for a circular aperture. Note that, similar to a single slit, the central maximum is wider and brighter than those to the sides. (b) Two point objects produce overlapping diffraction patterns. Shown here is the Rayleigh criterion for being just resolvable. The central maximum of one pattern lies on the first minimum of the other.

Connections: Limits to Knowledge

All attempts to observe the size and shape of objects are limited by the wavelength of the probe. Even the small wavelength of light prohibits exact precision. When extremely small wavelength probes as with an electron microscope are used, the system is disturbed, still limiting our knowledge, much as making an electrical measurement alters a circuit. Heisenberg’s uncertainty principle asserts that this limit is fundamental and inescapable, as we shall see in quantum mechanics.

Example 27.5 Calculating Diffraction Limits of the Hubble Space Telescope

The primary mirror of the orbiting Hubble Space Telescope has a diameter of 2.40 m. Being in orbit, this telescope avoids the degrading effects of atmospheric distortion on its resolution. (a) What is the angle between two just-resolvable point light sources (perhaps two stars)? Assume an average light wavelength of 550 nm. (b) If these two stars are at the 2 million light year distance of the Andromeda galaxy, how close together can they be and still be resolved? (A light year, or ly, is the distance light travels in 1 year.)

Strategy

The Rayleigh criterion stated in the equation \( \theta = \frac{1.22 \lambda}{D} \) gives the smallest possible angle \( \theta \) between point sources, or the best obtainable resolution. Once this angle is found, the distance between stars can be calculated, since we are given how far away they are.

Solution for (a)

The Rayleigh criterion for the minimum resolvable angle is

\[
\theta = \frac{1.22 \lambda}{D}.
\]

Entering known values gives

\[
\theta = \frac{1.22 \times 550 \times 10^{-9} \text{ m}}{2.40 \text{ m}} = 2.80 \times 10^{-7} \text{ rad}.
\]

Solution for (b)

The distance \( s \) between two objects a distance \( r \) away and separated by an angle \( \theta \) is \( s = r \theta \).

Substituting known values gives

\[
s = (2.0 \times 10^6 \text{ ly})(2.80 \times 10^{-7} \text{ rad}) = 0.56 \text{ ly}.
\]

Discussion

The angle found in part (a) is extraordinarily small (less than 1/50,000 of a degree), because the primary mirror is so large compared with the wavelength of light. As noticed, diffraction effects are most noticeable when light interacts with objects having sizes on the order of the wavelength of light. However, the effect is still there, and there is a diffraction limit to what is observable. The actual resolution of the Hubble Telescope is not quite as good as that found here. As with all instruments, there are other effects, such as non-uniformities in mirrors or aberrations in lenses that further limit resolution. However, Figure 27.27 gives an indication of the extent of the detail observable with the Hubble because of its size and quality and especially because it is above the Earth’s atmosphere.
Figure 27.27 These two photographs of the M82 galaxy give an idea of the observable detail using the Hubble Space Telescope compared with that using a ground-based telescope. (a) On the left is a ground-based image. (credit: Ricnun, Wikimedia Commons) (b) The photo on the right was captured by Hubble. (credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA))

The answer in part (b) indicates that two stars separated by about half a light year can be resolved. The average distance between stars in a galaxy is on the order of 5 light years in the outer parts and about 1 light year near the galactic center. Therefore, the Hubble can resolve most of the individual stars in Andromeda galaxy, even though it lies at such a huge distance that its light takes 2 million years for its light to reach us. Figure 27.28 shows another mirror used to observe radio waves from outer space.

Figure 27.28 A 305-m-diameter natural bowl at Arecibo in Puerto Rico is lined with reflective material, making it into a radio telescope. It is the largest curved focusing dish in the world. Although $D$ for Arecibo is much larger than for the Hubble Telescope, it detects much longer wavelength radiation and its diffraction limit is significantly poorer than Hubble's. Arecibo is still very useful, because important information is carried by radio waves that is not carried by visible light. (credit: Tatyana Temirbulatova, Flickr)

Diffraction is not only a problem for optical instruments but also for the electromagnetic radiation itself. Any beam of light having a finite diameter $D$ and a wavelength $\lambda$ exhibits diffraction spreading. The beam spreads out with an angle $\theta$ given by the equation $\theta = 1.22 \frac{\lambda}{D}$. Take, for example, a laser beam made of rays as parallel as possible (angles between rays as close to $\theta = 0^\circ$ as possible) instead spreads out at an angle $\theta = 1.22 \frac{\lambda}{D}$, where $D$ is the diameter of the beam and $\lambda$ is its wavelength. This spreading is impossible to observe for a flashlight, because its beam is not very parallel to start with. However, for long-distance transmission of laser beams or microwave signals, diffraction spreading can be significant (see Figure 27.29). To avoid this, we can increase $D$. This is done for laser light sent to the Moon to measure its distance from the Earth. The laser beam is expanded through a telescope to make $D$ much larger and $\theta$ smaller.

Figure 27.29 The beam produced by this microwave transmission antenna will spread out at a minimum angle $\theta = 1.22 \frac{\lambda}{D}$ due to diffraction. It is impossible to produce a near-parallel beam, because the beam has a limited diameter.

In most biology laboratories, resolution is presented when the use of the microscope is introduced. The ability of a lens to produce sharp images of two closely spaced point objects is called resolution. The smaller the distance $x$ by which two objects can be separated and still be seen as distinct, the greater the resolution. The resolving power of a lens is defined as that distance $x$. An expression for resolving power is obtained from the Rayleigh criterion. In Figure 27.30(a) we have two point objects separated by a distance $x$. According to the Rayleigh criterion, resolution is possible when the minimum angular separation is

$$\theta = 1.22 \frac{\lambda}{D} = \frac{x}{D}$$

(27.29)
where \( d \) is the distance between the specimen and the objective lens, and we have used the small angle approximation (i.e., we have assumed that \( x \) is much smaller than \( d \)), so that \( \tan \theta \approx \sin \theta \approx \theta \).

Therefore, the resolving power is

\[
x = 1.22 \frac{\lambda d}{D}.
\]

Another way to look at this is by re-examining the concept of Numerical Aperture (\( NA \)) discussed in Microscopes. There, \( NA \) is a measure of the maximum acceptance angle at which the fiber will take light and still contain it within the fiber. Figure 27.30(b) shows a lens and an object at point P. The \( NA \) here is a measure of the ability of the lens to gather light and resolve fine detail. The angle subtended by the lens at its focus is defined to be \( \theta = 2\alpha \). From the figure and again using the small angle approximation, we can write

\[
\sin \alpha = \frac{D/2}{d} = \frac{D}{2d}.
\]

The \( NA \) for a lens is \( NA = n \sin \alpha \), where \( n \) is the index of refraction of the medium between the objective lens and the object at point P.

From this definition for \( NA \), we can see that

\[
x = 1.22 \frac{\lambda d}{D} = 1.22 \frac{\lambda d}{2 \sin \alpha} = 0.61 \frac{\lambda n}{NA}.
\]

In a microscope, \( NA \) is important because it relates to the resolving power of a lens. A lens with a large \( NA \) will be able to resolve finer details. Lenses with larger \( NA \) will also be able to collect more light and so give a brighter image. Another way to describe this situation is that the larger the \( NA \), the larger the cone of light that can be brought into the lens, and so more of the diffraction modes will be collected. Thus the microscope has more information to form a clear image, and so its resolving power will be higher.

One of the consequences of diffraction is that the focal point of a beam has a finite width and intensity distribution. Consider focusing when only considering geometric optics, shown in Figure 27.31(a). The focal point is infinitely small with a huge intensity and the capacity to incinerate most samples irrespective of the \( NA \) of the objective lens. For wave optics, due to diffraction, the focal point spreads to become a focal spot (see Figure 27.31(b)) with the size of the spot decreasing with increasing \( NA \). Consequently, the intensity in the focal spot increases with increasing \( NA \). The higher the \( NA \), the greater the chances of photodegrading the specimen. However, the spot never becomes a true point.
27.7 Thin Film Interference

The bright colors seen in an oil slick floating on water or in a sunlit soap bubble are caused by interference. The brightest colors are those that interfere constructively. This interference is between light reflected from different surfaces of a thin film; thus, the effect is known as **thin film interference**. As noticed before, interference effects are most prominent when light interacts with something having a size similar to its wavelength. A thin film is one having a thickness $t$ smaller than a few times the wavelength of light, $\lambda$. Since color is associated indirectly with $\lambda$ and since all interference depends in some way on the ratio of $\lambda$ to the size of the object involved, we should expect to see different colors for different thicknesses of a film, as in **Figure 27.32**.

**Figure 27.32** These soap bubbles exhibit brilliant colors when exposed to sunlight. (credit: Scott Robinson, Flickr)

What causes thin film interference? **Figure 27.33** shows how light reflected from the top and bottom surfaces of a film can interfere. Incident light is only partially reflected from the top surface of the film (ray 1). The remainder enters the film and is itself partially reflected from the bottom surface. Part of the light reflected from the bottom surface can emerge from the top of the film (ray 2) and interfere with light reflected from the top (ray 1). Since the ray that enters the film travels a greater distance, it may be in or out of phase with the ray reflected from the top. However, consider for a moment, again, the bubbles in **Figure 27.32**. The bubbles are darkest where they are thinnest. Furthermore, if you observe a soap bubble carefully, you will note it gets dark at the point where it breaks. For very thin films, the difference in path lengths of ray 1 and ray 2 in **Figure 27.33** is negligible; so why should they interfere destructively and not constructively? The answer is that a phase change can occur upon reflection. The rule is as follows:

When light reflects from a medium having an index of refraction greater than that of the medium in which it is traveling, a 180° phase change (or a $\lambda/2$ shift) occurs.
If the film in Figure 27.33 is a soap bubble (essentially water with air on both sides), then there is a $\lambda/2$ shift for ray 1 and none for ray 2. Thus, when the film is very thin, the path length difference between the two rays is negligible, they are exactly out of phase, and destructive interference will occur at all wavelengths and so the soap bubble will be dark here.

The thickness of the film relative to the wavelength of light is the other crucial factor in thin film interference. Ray 2 in Figure 27.33 travels a greater distance than ray 1. For light incident perpendicular to the surface, ray 2 travels a distance approximately $2t$ farther than ray 1. When this distance is an integral or half-integral multiple of the wavelength in the medium ($\lambda_n = \lambda/n$, where $\lambda$ is the wavelength in vacuum and $n$ is the index of refraction), constructive or destructive interference occurs, depending also on whether there is a phase change in either ray.

**Example 27.6 Calculating Non-reflective Lens Coating Using Thin Film Interference**

Sophisticated cameras use a series of several lenses. Light can reflect from the surfaces of these various lenses and degrade image clarity. To limit these reflections, lenses are coated with a thin layer of magnesium fluoride that causes destructive thin film interference. What is the thinnest this film can be, if its index of refraction is 1.38 and it is designed to limit the reflection of 550-nm light, normally the most intense visible wavelength? The index of refraction of glass is 1.52.

**Strategy**

Refer to Figure 27.33 and use $n_1 = 100$ for air, $n_2 = 1.38$, and $n_3 = 1.52$. Both ray 1 and ray 2 will have a $\lambda/2$ shift upon reflection. Thus, to obtain destructive interference, ray 2 will need to travel a half wavelength farther than ray 1. For rays incident perpendicularly, the path length difference is $2t$.

**Solution**

To obtain destructive interference here,

$$2t = \frac{\lambda_{n_2}}{2},$$

(27.33)

where $\lambda_{n_2}$ is the wavelength in the film and is given by $\lambda_{n_2} = \frac{\lambda}{n_2}$.

Thus,

$$2t = \frac{\lambda}{n_2},$$

(27.34)

Solving for $t$ and entering known values yields

$$t = \frac{\lambda_{n_2}}{4} = \frac{(550 \text{ nm})/1.38}{4} = 99.6 \text{ nm}.$$  

(27.35)

**Discussion**
Films such as the one in this example are most effective in producing destructive interference when the thinnest layer is used, since light over a broader range of incident angles will be reduced in intensity. These films are called non-reflective coatings; this is only an approximately correct description, though, since other wavelengths will only be partially cancelled. Non-reflective coatings are used in car windows and sunglasses.

Thin film interference is most constructive or most destructive when the path length difference for the two rays is an integral or half-integral wavelength, respectively. That is, for rays incident perpendicularly, 
\[ 2t = \frac{\lambda}{n}, \frac{3\lambda}{n}, \frac{5\lambda}{n}, \ldots \text{ or } 2t = \frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}, \ldots . \]

To know whether interference is constructive or destructive, you must also determine if there is a phase change upon reflection. Thin film interference thus depends on film thickness, the wavelength of light, and the refractive indices. For white light incident on a film that varies in thickness, you will observe rainbow colors of constructive interference for various wavelengths as the thickness varies.

**Example 27.7 Soap Bubbles: More Than One Thickness can be Constructive**

(a) What are the three smallest thicknesses of a soap bubble that produce constructive interference for red light with a wavelength of 650 nm? The index of refraction of soap is taken to be the same as that of water. (b) What three smallest thicknesses will give destructive interference?

**Strategy and Concept**

Use Figure 27.33 to visualize the bubble. Note that \( n_1 = n_3 = 1.00 \) for air, and \( n_2 = 1.333 \) for soap (equivalent to water). There is a \( \lambda / 2 \) shift for ray 1 reflected from the top surface of the bubble, and no shift for ray 2 reflected from the bottom surface. To get constructive interference, then, the path length difference (\( 2t \)) must be a half-integral multiple of the wavelength—the first three being \( \frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4} \). To get destructive interference, the path length difference must be an integral multiple of the wavelength—the first three being 0, \( \frac{\lambda}{4} \), and \( \frac{3\lambda}{4} \).

**Solution for (a)**

**Constructive interference** occurs here when

\[ 2t_c = \frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4}, \ldots \]

The smallest constructive thickness \( t_c \) thus is

\[ t_c = \frac{\lambda}{4} = \frac{650 \text{ nm}}{1.333} \frac{4}{4} = 122 \text{ nm}. \]  

The next thickness that gives constructive interference is \( t'_c = \frac{3\lambda}{4} \), so that

\[ t'_c = 366 \text{ nm}. \]  

Finally, the third thickness producing constructive interference is \( t''_c = \frac{5\lambda}{4} \), so that

\[ t''_c = 610 \text{ nm}. \]

**Solution for (b)**

For **destructive interference**, the path length difference here is an integral multiple of the wavelength. The first occurs for zero thickness, since there is a phase change at the top surface. That is,

\[ t_d = 0. \]

The first non-zero thickness producing destructive interference is

\[ 2t'_d = \lambda. \]

Substituting known values gives

\[ t'_d = \frac{\lambda}{2} = \frac{650 \text{ nm}}{1.333} \frac{2}{2} = 244 \text{ nm}. \]  

Finally, the third destructive thickness is \( 2t''_d = 2\lambda \), so that

\[ t''_d = \lambda = \frac{650 \text{ nm}}{1.333} = 488 \text{ nm}. \]

**Discussion**

If the bubble was illuminated with pure red light, we would see bright and dark bands at very uniform increases in thickness. First would be a dark band at 0 thickness, then bright at 122 nm thickness, then dark at 244 nm, bright at 366 nm, dark at 488 nm, and bright at 610 nm. If the bubble varied smoothly in thickness, like a smooth wedge, then the bands would be evenly spaced.
Another example of thin film interference can be seen when microscope slides are separated (see Figure 27.34). The slides are very flat, so that the wedge of air between them increases in thickness very uniformly. A phase change occurs at the second surface but not the first, and so there is a dark band where the slides touch. The rainbow colors of constructive interference repeat, going from violet to red again and again as the distance between the slides increases. As the layer of air increases, the bands become more difficult to see, because slight changes in incident angle have greater effects on path length differences. If pure-wavelength light instead of white light is used, then bright and dark bands are obtained rather than repeating rainbow colors.

![Figure 27.34](image)

Figure 27.34 (a) The rainbow color bands are produced by thin film interference in the air between the two glass slides. (b) Schematic of the paths taken by rays in the wedge of air between the slides.

An important application of thin film interference is found in the manufacturing of optical instruments. A lens or mirror can be compared with a master as it is being ground, allowing it to be shaped to an accuracy of less than a wavelength over its entire surface. Figure 27.35 illustrates the phenomenon called Newton's rings, which occurs when the plane surfaces of two lenses are placed together. (The circular bands are called Newton's rings because Isaac Newton described them and their use in detail. Newton did not discover them; Robert Hooke did, and Newton did not believe they were due to the wave character of light.) Each successive ring of a given color indicates an increase of only one wavelength in the distance between the lens and the blank, so that great precision can be obtained. Once the lens is perfect, there will be no rings.

![Figure 27.35](image)

Figure 27.35 “Newton’s rings” interference fringes are produced when two plano-convex lenses are placed together with their plane surfaces in contact. The rings are created by interference between the light reflected off the two surfaces as a result of a slight gap between them, indicating that these surfaces are not precisely plane but are slightly convex. (credit: Ulf Seifert, Wikimedia Commons)

The wings of certain moths and butterflies have nearly iridescent colors due to thin film interference. In addition to pigmentation, the wing’s color is affected greatly by constructive interference of certain wavelengths reflected from its film-coated surface. Car manufacturers are offering special paint jobs that use thin film interference to produce colors that change with angle. This expensive option is based on variation of thin film path length differences with angle. Security features on credit cards, banknotes, driving licenses and similar items prone to forgery use thin film interference, diffraction gratings, or holograms. Australia led the way with dollar bills printed on polymer with a diffraction grating security feature making the currency difficult to forge. Other countries such as New Zealand and Taiwan are using similar technologies, while the United States currency includes a thin film interference effect.

**Making Connections: Take-Home Experiment—Thin Film Interference**

One feature of thin film interference and diffraction gratings is that the pattern shifts as you change the angle at which you look or move your head. Find examples of thin film interference and gratings around you. Explain how the patterns change for each specific example. Find examples where the thickness changes giving rise to changing colors. If you can find two microscope slides, then try observing the effect shown in Figure 27.34. Try separating one end of the two slides with a hair or maybe a thin piece of paper and observe the effect.

**Problem-Solving Strategies for Wave Optics**

**Step 1.** **Examine the situation to determine that interference is involved.** Identify whether slits or thin film interference are considered in the problem.

**Step 2.** **If slits are involved,** note that diffraction gratings and double slits produce very similar interference patterns, but that gratings have narrower (sharper) maxima. Single slit patterns are characterized by a large central maximum and smaller maxima to the sides.

**Step 3.** **If thin film interference is involved,** take note of the path length difference between the two rays that interfere. Be certain to use the wavelength in the medium involved, since it differs from the wavelength in vacuum. Note also that there is an additional $\lambda/2$ phase shift when light reflects from a medium with a greater index of refraction.

**Step 4.** **Identify exactly what needs to be determined in the problem (identify the unknowns).** A written list is useful. Draw a diagram of the situation. Labeling the diagram is useful.

**Step 5.** **Make a list of what is given or can be inferred from the problem as stated (identify the knowns).**
**Step 6.** Solve the appropriate equation for the quantity to be determined (the unknown), and enter the knowns. Slits, gratings, and the Rayleigh limit involve equations.

**Step 7.** For thin film interference, you will have constructive interference for a total shift that is an integral number of wavelengths. You will have destructive interference for a total shift of a half-integral number of wavelengths. Always keep in mind that crest to crest is constructive whereas crest to trough is destructive.

**Step 8.** Check to see if the answer is reasonable: Does it make sense? Angles in interference patterns cannot be greater than 90º, for example.

### 27.8 Polarization

Polaroid sunglasses are familiar to most of us. They have a special ability to cut the glare of light reflected from water or glass (see Figure 27.36). Polaroids have this ability because of a wave characteristic of light called polarization. What is polarization? How is it produced? What are some of its uses? The answers to these questions are related to the wave character of light.

![Figure 27.36](https://example.com/figure2736.png)

*Figure 27.36* These two photographs of a river show the effect of a polarizing filter in reducing glare in light reflected from the surface of water. Part (b) of this figure was taken with a polarizing filter and part (a) was not. As a result, the reflection of clouds and sky observed in part (a) is not observed in part (b). Polarizing sunglasses are particularly useful on snow and water. (credit: Amithshs, Wikimedia Commons)

Light is one type of electromagnetic (EM) wave. As noted earlier, EM waves are transverse waves consisting of varying electric and magnetic fields that oscillate perpendicular to the direction of propagation (see Figure 27.37). There are specific directions for the oscillations of the electric and magnetic fields. **Polarization** is the attribute that a wave’s oscillations have a definite direction relative to the direction of propagation of the wave. (This is not the same type of polarization as that discussed for the separation of charges.) Waves having such a direction are said to be **polarized**. For an EM wave, we define the **direction of polarization** to be the direction parallel to the electric field. Thus we can think of the electric field arrows as showing the direction of polarization, as in Figure 27.37.

![Figure 27.37](https://example.com/figure2737.png)

*Figure 27.37* An EM wave, such as light, is a transverse wave. The electric and magnetic fields are perpendicular to the direction of propagation.

To examine this further, consider the transverse waves in the ropes shown in Figure 27.38. The oscillations in one rope are in a vertical plane and are said to be **vertically polarized**. Those in the other rope are in a horizontal plane and are **horizontally polarized**. If a vertical slit is placed on the first rope, the waves pass through. However, a vertical slit blocks the horizontally polarized waves. For EM waves, the direction of the electric field is analogous to the disturbances on the ropes.

![Figure 27.38](https://example.com/figure2738.png)

*Figure 27.38* The transverse oscillations in one rope are in a vertical plane, and those in the other rope are in a horizontal plane. The first is said to be vertically polarized, and the other is said to be horizontally polarized. Vertical slits pass vertically polarized waves and block horizontally polarized waves.

The Sun and many other light sources produce waves that are randomly polarized (see Figure 27.39). Such light is said to be **unpolarized** because it is composed of many waves with all possible directions of polarization. Polaroid materials, invented by the founder of Polaroid Corporation, Edwin
Land, act as a polarizing slit for light, allowing only polarization in one direction to pass through. Polarizing filters are composed of long molecules aligned in one direction. Thinking of the molecules as many slits, analogous to those for the oscillating ropes, we can understand why only light with a specific polarization can get through. The **axis of a polarizing filter** is the direction along which the filter passes the electric field of an EM wave (see Figure 27.40).

![Random polarization](image)

*Figure 27.39* The slender arrow represents a ray of unpolarized light. The bold arrows represent the direction of polarization of the individual waves composing the ray. Since the light is unpolarized, the arrows point in all directions.

![Polarizing filter](image)

*Figure 27.40* A polarizing filter has a polarization axis that acts as a slit passing through electric fields parallel to its direction. The direction of polarization of an EM wave is defined to be the direction of its electric field.

*Figure 27.41* shows the effect of two polarizing filters on originally unpolarized light. The first filter polarizes the light along its axis. When the axes of the first and second filters are aligned (parallel), then all of the polarized light passed by the first filter is also passed by the second. If the second polarizing filter is rotated, only the component of the light parallel to the second filter’s axis is passed. When the axes are perpendicular, no light is passed by the second.

Only the component of the EM wave parallel to the axis of a filter is passed. Let us call the angle between the direction of polarization and the axis of a filter \( \theta \). If the electric field has an amplitude \( E \), then the transmitted part of the wave has an amplitude \( E \cos \theta \) (see Figure 27.42). Since the intensity of a wave is proportional to its amplitude squared, the intensity \( I \) of the transmitted wave is related to the incident wave by

\[
I = I_0 \cos^2 \theta, \tag{27.44}
\]

where \( I_0 \) is the intensity of the polarized wave before passing through the filter. (The above equation is known as Malus’s law.)
**Example 27.8 Calculating Intensity Reduction by a Polarizing Filter**

What angle is needed between the direction of polarized light and the axis of a polarizing filter to reduce its intensity by 90.0%?

**Strategy**

When the intensity is reduced by 90.0%, it is 10.0% or 0.100 times its original value. That is, \( I = 0.100I_0 \). Using this information, the equation \( I = I_0 \cos^2 \theta \) can be used to solve for the needed angle.

**Solution**

Solving the equation \( I = I_0 \cos^2 \theta \) for \( \cos \theta \) and substituting with the relationship between \( I \) and \( I_0 \) gives

\[
\cos \theta = \frac{I}{I_0} = \frac{0.100I_0}{I_0} = 0.3162.
\]

Solving for \( \theta \) yields

\[
\theta = \cos^{-1} 0.3162 = 71.6^\circ.
\]

**Discussion**

A fairly large angle between the direction of polarization and the filter axis is needed to reduce the intensity to 10.0% of its original value. This seems reasonable based on experimenting with polarizing films. It is interesting that, at an angle of 45°, the intensity is reduced to 50% of its original value (as you will show in this section’s Problems & Exercises). Note that 71.6° is 18.4° from reducing the intensity to zero, and that at an angle of 18.4° the intensity is reduced to 90.0% of its original value (as you will also show in Problems & Exercises), giving evidence of symmetry.

**Polarization by Reflection**

By now you can probably guess that Polaroid sunglasses cut the glare in reflected light because that light is polarized. You can check this for yourself by holding Polaroid sunglasses in front of you and rotating them while looking at light reflected from water or glass. As you rotate the sunglasses, you...
will notice the light gets bright and dim, but not completely black. This implies the reflected light is partially polarized and cannot be completely blocked by a polarizing filter.

**Figure 27.43** illustrates what happens when unpolarized light is reflected from a surface. Vertically polarized light is preferentially refracted at the surface, so that the reflected light is left more horizontally polarized. The reasons for this phenomenon are beyond the scope of this text, but a convenient mnemonic for remembering this is to imagine the polarization direction to be like an arrow. Vertical polarization would be like an arrow perpendicular to the surface and would be more likely to stick and not be reflected. Horizontal polarization is like an arrow bouncing on its side and would be more likely to be reflected. Sunglasses with vertical axes would then block more reflected light than unpolarized light from other sources.

Since the part of the light that is not reflected is refracted, the amount of polarization depends on the indices of refraction of the media involved. It can be shown that reflected light is completely polarized at a angle of reflection \( \theta_b \), given by

\[
\tan \theta_b = \frac{n_2}{n_1},
\]

where \( n_1 \) is the medium in which the incident and reflected light travel and \( n_2 \) is the index of refraction of the medium that forms the interface that reflects the light. This equation is known as Brewster’s law, and \( \theta_b \) is known as Brewster’s angle, named after the 19th-century Scottish physicist who discovered them.

**Things Great and Small: Atomic Explanation of Polarizing Filters**

Polarizing filters have a polarization axis that acts as a slit. This slit passes electromagnetic waves (often visible light) that have an electric field parallel to the axis. This is accomplished with long molecules aligned perpendicular to the axis as shown in **Figure 27.44**.

**Figure 27.44** Long molecules are aligned perpendicular to the axis of a polarizing filter. The component of the electric field in an EM wave perpendicular to these molecules passes through the filter, while the component parallel to the molecules is absorbed.

**Figure 27.45** illustrates how the component of the electric field parallel to the long molecules is absorbed. An electromagnetic wave is composed of oscillating electric and magnetic fields. The electric field is strong compared with the magnetic field and is more effective in exerting force on charges in the molecules. The most affected charged particles are the electrons in the molecules, since electron masses are small. If the electron is forced to oscillate, it can absorb energy from the EM wave. This reduces the fields in the wave and, hence, reduces its intensity. In long molecules, electrons can more easily oscillate parallel to the molecule than in the perpendicular direction. The electrons are bound to the
molecule and are more restricted in their movement perpendicular to the molecule. Thus, the electrons can absorb EM waves that have a component of their electric field parallel to the molecule. The electrons are much less responsive to electric fields perpendicular to the molecule and will allow those fields to pass. Thus the axis of the polarizing filter is perpendicular to the length of the molecule.

Figure 27.45 Artist’s conception of an electron in a long molecule oscillating parallel to the molecule. The oscillation of the electron absorbs energy and reduces the intensity of the component of the EM wave that is parallel to the molecule.

Example 27.9 Calculating Polarization by Reflection

(a) At what angle will light traveling in air be completely polarized horizontally when reflected from water? (b) From glass?

Strategy
All we need to solve these problems are the indices of refraction. Air has \( n_1 = 1.00 \), water has \( n_2 = 1.333 \), and crown glass has \( n'_2 = 1.520 \). The equation \( \tan \theta_b = \frac{n_2}{n_1} \) can be directly applied to find \( \theta_b \) in each case.

Solution for (a)
Putting the known quantities into the equation

\[
\tan \theta_b = \frac{n_2}{n_1}
\]

(27.48)

gives

\[
\tan \theta_b = \frac{n_2}{n_1} = \frac{1.333}{1.00} = 1.333.
\]

(27.49)

Solving for the angle \( \theta_b \) yields

\[
\theta_b = \tan^{-1} 1.333 = 53.1^\circ.
\]

(27.50)

Solution for (b)
Similarly, for crown glass and air,

\[
\tan \theta'_b = \frac{n'_2}{n_1} = \frac{1.520}{1.00} = 1.52.
\]

(27.51)

Thus,

\[
\theta'_b = \tan^{-1} 1.52 = 56.7^\circ.
\]

(27.52)

Discussion
Light reflected at these angles could be completely blocked by a good polarizing filter held with its axis vertical. Brewster’s angle for water and air are similar to those for glass and air, so that sunglasses are equally effective for light reflected from either water or glass under similar circumstances. Light not reflected is refracted into these media. So at an incident angle equal to Brewster’s angle, the refracted light will be
slightly polarized vertically. It will not be completely polarized vertically, because only a small fraction of the incident light is reflected, and so a significant amount of horizontally polarized light is refracted.

**Polarization by Scattering**

If you hold your Polaroid sunglasses in front of you and rotate them while looking at blue sky, you will see the sky get bright and dim. This is a clear indication that light scattered by air is partially polarized. Figure 27.46 helps illustrate how this happens. Since light is a transverse EM wave, it vibrates the electrons of air molecules perpendicular to the direction it is traveling. The electrons then radiate like small antennae. Since they are oscillating perpendicular to the direction of the light ray, they produce EM radiation that is polarized perpendicular to the direction of the ray. When viewing the light along a line perpendicular to the original ray, as in Figure 27.46, there can be no polarization in the scattered light parallel to the original ray, because that would require the original ray to be a longitudinal wave. Along other directions, a component of the other polarization can be projected along the line of sight, and the scattered light will only be partially polarized. Furthermore, multiple scattering can bring light to your eyes from other directions and can contain different polarizations.

![Figure 27.46](image)

**Photographs of the sky can be darkened by polarizing filters, a trick used by many photographers to make clouds brighter by contrast.** Scattering from other particles, such as smoke or dust, can also polarize light. Detecting polarization in scattered EM waves can be a useful analytical tool in determining the scattering source.

There is a range of optical effects used in sunglasses. Besides being Polaroid, other sunglasses have colored pigments embedded in them, while others use non-reflective or even reflective coatings. A recent development is photochromic lenses, which darken in the sunlight and become clear indoors. Photochromic lenses are embedded with organic microcrystalline molecules that change their properties when exposed to UV in sunlight, but become clear in artificial lighting with no UV.

**Take-Home Experiment: Polarization**

Find Polaroid sunglasses and rotate one while holding the other still and look at different surfaces and objects. Explain your observations. What is the difference in angle from when you see a maximum intensity to when you see a minimum intensity? Find a reflective glass surface and do the same. At what angle does the glass need to be oriented to give minimum glare?

**Liquid Crystals and Other Polarization Effects in Materials**

While you are undoubtedly aware of liquid crystal displays (LCDs) found in watches, calculators, computer screens, cellphones, flat screen televisions, and other myriad places, you may not be aware that they are based on polarization. Liquid crystals are so named because their molecules can be aligned even though they are in a liquid. Liquid crystals have the property that they can rotate the polarization of light passing through them by $90^\circ$. Furthermore, this property can be turned off by the application of a voltage, as illustrated in Figure 27.47. It is possible to manipulate this characteristic quickly and in small well-defined regions to create the contrast patterns we see in so many LCD devices.

In flat screen LCD televisions, there is a large light at the back of the TV. The light travels to the front screen through millions of tiny units called pixels (picture elements). One of these is shown in Figure 27.47 (a) and (b). Each unit has three cells, with red, blue, or green filters, each controlled independently. When the voltage across a liquid crystal is switched off, the liquid crystal passes the light through the particular filter. One can vary the picture contrast by varying the strength of the voltage applied to the liquid crystal.
Many crystals and solutions rotate the plane of polarization of light passing through them. Such substances are said to be **optically active**. Examples include sugar water, insulin, and collagen (see **Figure 27.48**). In addition to depending on the type of substance, the amount and direction of rotation depends on a number of factors. Among these is the concentration of the substance, the distance the light travels through it, and the wavelength of light. Optical activity is due to the asymmetric shape of molecules in the substance, such as being helical. Measurements of the rotation of polarized light passing through substances can thus be used to measure concentrations, a standard technique for sugars. It can also give information on the shapes of molecules, such as proteins, and factors that affect their shapes, such as temperature and pH.

**Figure 27.48** Optical activity is the ability of some substances to rotate the plane of polarization of light passing through them. The rotation is detected with a polarizing filter or analyzer.

Glass and plastic become optically active when stressed; the greater the stress, the greater the effect. Optical stress analysis on complicated shapes can be performed by making plastic models of them and observing them through crossed filters, as seen in **Figure 27.49**. It is apparent that the effect depends on wavelength as well as stress. The wavelength dependence is sometimes also used for artistic purposes.
Another interesting phenomenon associated with polarized light is the ability of some crystals to split an unpolarized beam of light into two. Such crystals are said to be birefringent (see Figure 27.50). Each of the separated rays has a specific polarization. One behaves normally and is called the ordinary ray, whereas the other does not obey Snell’s law and is called the extraordinary ray. Birefringent crystals can be used to produce polarized beams from unpolarized light. Some birefringent materials preferentially absorb one of the polarizations. These materials are called dichroic and can produce polarization by this preferential absorption. This is fundamentally how polarizing filters and other polarizers work. The interested reader is invited to further pursue the numerous properties of materials related to polarization.

Figure 27.50 Birefringent materials, such as the common mineral calcite, split unpolarized beams of light into two. The ordinary ray behaves as expected, but the extraordinary ray does not obey Snell’s law.

27.9 *Extended Topic* Microscopy Enhanced by the Wave Characteristics of Light

Physics research underpins the advancement of developments in microscopy. As we gain knowledge of the wave nature of electromagnetic waves and methods to analyze and interpret signals, new microscopes that enable us to “see” more are being developed. It is the evolution and newer generation of microscopes that are described in this section.

The use of microscopes (microscopy) to observe small details is limited by the wave nature of light. Owing to the fact that light diffracts significantly around small objects, it becomes impossible to observe details significantly smaller than the wavelength of light. One rule of thumb has it that all details smaller than about \( \lambda \) are difficult to observe. Radar, for example, can detect the size of an aircraft, but not its individual rivets, since the wavelength of most radar is several centimeters or greater. Similarly, visible light cannot detect individual atoms, since atoms are about 0.1 nm in size and visible wavelengths range from 380 to 760 nm. Ironically, special techniques used to obtain the best possible resolution with microscopes take advantage of the same wave characteristics of light that ultimately limit the detail.

The most obvious method of obtaining better detail is to utilize shorter wavelengths. Ultraviolet (UV) microscopes have been constructed with special lenses that transmit UV rays and utilize photographic or electronic techniques to record images. The shorter UV wavelengths allow somewhat greater detail to be observed, but drawbacks, such as the hazard of UV to living tissue and the need for special detection devices and lenses (which tend to be dispersive in the UV), severely limit the use of UV microscopes. Elsewhere, we will explore practical uses of very short wavelength EM waves, such as x rays, and other short-wavelength probes, such as electrons in electron microscopes, to detect small details.

Another difficulty in microscopy is the fact that many microscopic objects do not absorb much of the light passing through them. Contrast is the difference in intensity between objects and the background on which they are observed. Stains (such as dyes, fluorophores, etc.) are commonly employed to enhance contrast, but these tend to be application specific. More general wave interference techniques can be used to produce contrast. Figure 27.51 shows the passage of light through a sample. Since the indices of refraction differ, the number of wavelengths in the paths differs. Light emerging from the object is thus out of phase with light from the background and will interfere differently, producing enhanced contrast, especially if the light is coherent and monochromatic—as in laser light.
Figure 27.51 Light rays passing through a sample under a microscope will emerge with different phases depending on their paths. The object shown has a greater index of refraction than the background, and so the wavelength decreases as the ray passes through it. Superimposing these rays produces interference that varies with path, enhancing contrast between the object and background.

Interference microscopes enhance contrast between objects and background by superimposing a reference beam of light upon the light emerging from the sample. Since light from the background and objects differ in phase, there will be different amounts of constructive and destructive interference, producing the desired contrast in final intensity. Figure 27.52 shows schematically how this is done. Parallel rays of light from a source are split into two beams by a half-silvered mirror. These beams are called the object and reference beams. Each beam passes through identical optical elements, except that the object beam passes through the object we wish to observe microscopically. The light beams are recombined by another half-silvered mirror and interfere. Since the light rays passing through different parts of the object have different phases, interference will be significantly different and, hence, have greater contrast between them.

Another type of microscope utilizing wave interference and differences in phases to enhance contrast is called the phase-contrast microscope. While its principle is the same as the interference microscope, the phase-contrast microscope is simpler to use and construct. Its impact (and the principle upon which it is based) was so important that its developer, the Dutch physicist Frits Zernike (1888–1966), was awarded the Nobel Prize in 1953. Figure 27.53 shows the basic construction of a phase-contrast microscope. Phase differences between light passing through the object and background are produced by passing the rays through different parts of a phase plate (so called because it shifts the phase of the light passing through it). These two light rays are superimposed in the image plane, producing contrast due to their interference.
A polarization microscope also enhances contrast by utilizing a wave characteristic of light. Polarization microscopes are useful for objects that are optically active or birefringent, particularly if those characteristics vary from place to place in the object. Polarized light is sent through the object and then observed through a polarizing filter that is perpendicular to the original polarization direction. Nearly transparent objects can then appear with strong color and in high contrast. Many polarization effects are wavelength dependent, producing color in the processed image. Contrast results from the action of the polarizing filter in passing only components parallel to its axis.

Apart from the UV microscope, the variations of microscopy discussed so far in this section are available as attachments to fairly standard microscopes or as slight variations. The next level of sophistication is provided by commercial confocal microscopes, which use the extended focal region shown in Figure 27.31(b) to obtain three-dimensional images rather than two-dimensional images. Here, only a single plane or region of focus is identified; out-of-focus regions above and below this plane are subtracted out by a computer so the image quality is much better. This type of microscope makes use of fluorescence, where a laser provides the excitation light. Laser light passing through a tiny aperture called a pinhole forms an extended focal region within the specimen. The reflected light passes through the objective lens to a second pinhole and the photomultiplier detector, see Figure 27.54. The second pinhole is the key here and serves to block much of the light from points that are not at the focal point of the objective lens. The pinhole is conjugate (coupled) to the focal point of the lens. The second pinhole and detector are scanned, allowing reflected light from a small region or section of the extended focal region to be imaged at any one time. The out-of-focus light is excluded. Each image is stored in a computer, and a full scanned image is generated in a short time. Live cell processes can also be imaged at adequate scanning speeds allowing the imaging of three-dimensional microscopic movement. Confocal microscopy enhances images over conventional optical microscopy, especially for thicker specimens, and so has become quite popular.

The next level of sophistication is provided by microscopes attached to instruments that isolate and detect only a small wavelength band of light—monochromators and spectral analyzers. Here, the monochromatic light from a laser is scattered from the specimen. This scattered light shifts up or down as it excites particular energy levels in the sample. The uniqueness of the observed scattered light can give detailed information about the chemical composition of a given spot on the sample with high contrast—like molecular fingerprints. Applications are in materials science, nanotechnology, and the biomedical field. Fine details in biochemical processes over time can even be detected. The ultimate in microscopy is the electron microscope—to be discussed later. Research is being conducted into the development of new prototype microscopes that can become commercially available, providing better diagnostic and research capacities.
Figure 27.54 A confocal microscope provides three-dimensional images using pinholes and the extended depth of focus as described by wave optics. The right pinhole illuminates a tiny region of the sample in the focal plane. In-focus light rays from this tiny region pass through the dichroic mirror and the second pinhole to a detector and a computer. Out-of-focus light rays are blocked. The pinhole is scanned sideways to form an image of the entire focal plane. The pinhole can then be scanned up and down to gather images from different focal planes. The result is a three-dimensional image of the specimen.

Glossary

axis of a polarizing filter: the direction along which the filter passes the electric field of an EM wave

Brewster’s angle: \( \theta_b = \tan^{-1}\left(\frac{n_2}{n_1}\right) \) where \( n_2 \) is the index of refraction of the medium from which the light is reflected and \( n_1 \) is the index of refraction of the medium in which the reflected light travels

Brewster’s law: \( \tan \theta_b = \frac{n_2}{n_1} \), where \( n_1 \) is the medium in which the incident and reflected light travel and \( n_2 \) is the index of refraction of the medium that forms the interface that reflects the light

birefringent: crystals that split an unpolarized beam of light into two beams

cohherent: waves are in phase or have a definite phase relationship

confocal microscopes: microscopes that use the extended focal region to obtain three-dimensional images rather than two-dimensional images

constructive interference for a diffraction grating: occurs when the condition \( d \sin \theta = m\lambda \) (for \( m = 0, 1, -1, 2, -2, \ldots \)) is satisfied, where \( d \) is the distance between slits in the grating, \( \lambda \) is the wavelength of light, and \( m \) is the order of the maximum

constructive interference for a double slit: the path length difference must be an integral multiple of the wavelength

contrast: the difference in intensity between objects and the background on which they are observed

destructive interference for a double slit: the path length difference must be a half-integral multiple of the wavelength

destructive interference for a single slit: occurs when \( D \sin \theta = m\lambda \), (for \( m = 1, -1, 2, -2, 3, \ldots \)), where \( D \) is the slit width, \( \lambda \) is the light’s wavelength, \( \theta \) is the angle relative to the original direction of the light, and \( m \) is the order of the minimum

diffraction grating: a large number of evenly spaced parallel slits

diffraction: the bending of a wave around the edges of an opening or an obstacle

direction of polarization: the direction parallel to the electric field for EM waves

Huygens’s principle: every point on a wavefront is a source of wavelets that spread out in the forward direction at the same speed as the wave itself. The new wavefront is a line tangent to all of the wavelets

horizontally polarized: the oscillations are in a horizontal plane

incoherent: waves have random phase relationships

interference microscopes: microscopes that enhance contrast between objects and background by superimposing a reference beam of light upon the light emerging from the sample

optically active: substances that rotate the plane of polarization of light passing through them

order: the integer \( m \) used in the equations for constructive and destructive interference for a double slit
**phase-contrast microscope**: microscope utilizing wave interference and differences in phases to enhance contrast

**polarization microscope**: microscope that enhances contrast by utilizing a wave characteristic of light, useful for objects that are optically active

**polarization**: the attribute that wave oscillations have a definite direction relative to the direction of propagation of the wave

**polarized**: waves having the electric and magnetic field oscillations in a definite direction

**Rayleigh criterion**: two images are just resolvable when the center of the diffraction pattern of one is directly over the first minimum of the diffraction pattern of the other

**reflected light that is completely polarized**: light reflected at the angle of reflection $\theta_b$, known as Brewster's angle

**thin film interference**: interference between light reflected from different surfaces of a thin film

**ultraviolet (UV) microscopes**: microscopes constructed with special lenses that transmit UV rays and utilize photographic or electronic techniques to record images

**unpolarized**: waves that are randomly polarized

**vertically polarized**: the oscillations are in a vertical plane

**wavelength in a medium**: $\lambda_n = \lambda / n$, where $\lambda$ is the wavelength in vacuum, and $n$ is the index of refraction of the medium

---

**Section Summary**

27.1 The Wave Aspect of Light: Interference
- Wave optics is the branch of optics that must be used when light interacts with small objects or whenever the wave characteristics of light are considered.
- Wave characteristics are those associated with interference and diffraction.
- Visible light is the type of electromagnetic wave to which our eyes respond and has a wavelength in the range of 380 to 760 nm.
- Like all EM waves, the following relationship is valid in vacuum: $c = f \lambda$, where $c = 3 \times 10^8$ m/s is the speed of light, $f$ is the frequency of the electromagnetic wave, and $\lambda$ is its wavelength in vacuum.
- The wavelength $\lambda_n$ of light in a medium with index of refraction $n$ is $\lambda_n = \lambda / n$. Its frequency is the same as in vacuum.

27.2 Huygens’s Principle: Diffraction
- An accurate technique for determining how and where waves propagate is given by Huygens’s principle: Every point on a wavefront is a source of wavelets that spread out in the forward direction at the same speed as the wave itself. The new wavefront is a line tangent to all of the wavelets.
- Diffraction is the bending of a wave around the edges of an opening or other obstacle.

27.3 Young’s Double Slit Experiment
- Young’s double slit experiment gave definitive proof of the wave character of light.
- An interference pattern is obtained by the superposition of light from two slits.
- There is constructive interference when $d \sin \theta = m\lambda$ (for $m = 0, 1, -1, 2, -2, \ldots$), where $d$ is the distance between the slits, $\theta$ is the angle relative to the incident direction, and $m$ is the order of the interference.
- There is destructive interference when $d \sin \theta = \left(m + \frac{1}{2}\right)\lambda$ (for $m = 0, 1, -1, 2, -2, \ldots$).

27.4 Multiple Slit Diffraction
- A diffraction grating is a large collection of evenly spaced parallel slits that produces an interference pattern similar to but sharper than that of a double slit.
- There is constructive interference for a diffraction grating when $d \sin \theta = m\lambda$ (for $m = 0, 1, -1, 2, -2, \ldots$), where $d$ is the distance between slits in the grating, $\lambda$ is the wavelength of light, and $m$ is the order of the maximum.

27.5 Single Slit Diffraction
- A single slit produces an interference pattern characterized by a broad central maximum with narrower and dimmer maxima to the sides.
- There is destructive interference for a single slit when $D \sin \theta = m\lambda$, (for $m = 1, -1, 2, -2, 3, \ldots$), where $D$ is the slit width, $\lambda$ is the light’s wavelength, $\theta$ is the angle relative to the original direction of the light, and $m$ is the order of the minimum. Note that there is no $m = 0$ minimum.

27.6 Limits of Resolution: The Rayleigh Criterion
- Diffraction limits resolution.
- For a circular aperture, lens, or mirror, the Rayleigh criterion states that two images are just resolvable when the center of the diffraction pattern of one is directly over the first minimum of the diffraction pattern of the other.
• This occurs for two point objects separated by the angle \( \theta = 1.22 \frac{\lambda}{D} \), where \( \lambda \) is the wavelength of light (or other electromagnetic radiation) and \( D \) is the diameter of the aperture, lens, mirror, etc. This equation also gives the angular spreading of a source of light having a diameter \( D \).

27.7 Thin Film Interference
• Thin film interference occurs between the light reflected from the top and bottom surfaces of a film. In addition to the path length difference, there can be a phase change.
• When light reflects from a medium having an index of refraction greater than that of the medium in which it is traveling, a 180° phase change (or a \( \frac{\lambda}{2} \) shift) occurs.

27.8 Polarization
• Polarization is the attribute that wave oscillations have a definite direction relative to the direction of propagation of the wave.
• EM waves are transverse waves that may be polarized.
• The direction of polarization is defined to be the direction parallel to the electric field of the EM wave.
• Unpolarized light is composed of many rays having random polarization directions.
• Light can be polarized by passing it through a polarizing filter or other polarizing material. The intensity \( I \) of polarized light after passing through a polarizing filter is \( I = I_0 \cos^2 \theta \), where \( I_0 \) is the original intensity and \( \theta \) is the angle between the direction of polarization and the axis of the filter.
• Polarization is also produced by reflection.
• Brewster's law states that reflected light will be completely polarized at the angle of reflection \( \theta_b \), known as Brewster’s angle, given by a statement known as Brewster’s law: \( \tan \theta_b = \frac{n_2}{n_1} \), where \( n_1 \) is the medium in which the incident and reflected light travel and \( n_2 \) is the index of refraction of the medium that forms the interface that reflects the light.
• Polarization can also be produced by scattering.
• There are a number of types of optically active substances that rotate the direction of polarization of light passing through them.

27.9 *Extended Topic* Microscopy Enhanced by the Wave Characteristics of Light
• To improve microscope images, various techniques utilizing the wave characteristics of light have been developed. Many of these enhance contrast with interference effects.

Conceptual Questions

27.1 The Wave Aspect of Light: Interference
1. What type of experimental evidence indicates that light is a wave?
2. Give an example of a wave characteristic of light that is easily observed outside the laboratory.

27.2 Huygens’s Principle: Diffraction
3. How do wave effects depend on the size of the object with which the wave interacts? For example, why does sound bend around the corner of a building while light does not?
4. Under what conditions can light be modeled like a ray? Like a wave?
5. Go outside in the sunlight and observe your shadow. It has fuzzy edges even if you do not. Is this a diffraction effect? Explain.
6. Why does the wavelength of light decrease when it passes from vacuum into a medium? State which attributes change and which stay the same and, thus, require the wavelength to decrease.
7. Does Huygens’s principle apply to all types of waves?

27.3 Young’s Double Slit Experiment
8. Young’s double slit experiment breaks a single light beam into two sources. Would the same pattern be obtained for two independent sources of light, such as the headlights of a distant car? Explain.
9. Suppose you use the same double slit to perform Young’s double slit experiment in air and then repeat the experiment in water. Do the angles to the same parts of the interference pattern get larger or smaller? Does the color of the light change? Explain.
10. Is it possible to create a situation in which there is only destructive interference? Explain.
11. Figure 27.55 shows the central part of the interference pattern for a pure wavelength of red light projected onto a double slit. The pattern is actually a combination of single slit and double slit interference. Note that the bright spots are evenly spaced. Is this a double slit or single slit characteristic? Note that some of the bright spots are dim on either side of the center. Is this a single slit or double slit characteristic? Which is smaller, the slit width or the separation between slits? Explain your responses.

Figure 27.55 This double slit interference pattern also shows signs of single slit interference. (credit: PASCO)
27.4 Multiple Slit Diffraction
12. What is the advantage of a diffraction grating over a double slit in dispersing light into a spectrum?
13. What are the advantages of a diffraction grating over a prism in dispersing light for spectral analysis?
14. Can the lines in a diffraction grating be too close together to be useful as a spectroscopic tool for visible light? If so, what type of EM radiation would the grating be suitable for? Explain.
15. If a beam of white light passes through a diffraction grating with vertical lines, the light is dispersed into rainbow colors on the right and left. If a glass prism disperses white light to the right into a rainbow, how does the sequence of colors compare with that produced on the right by a diffraction grating?
16. Suppose pure-wavelength light falls on a diffraction grating. What happens to the interference pattern if the same light falls on a grating that has more lines per centimeter? What happens to the interference pattern if a longer-wavelength light falls on the same grating? Explain how these two effects are consistent in terms of the relationship of wavelength to the distance between slits.
17. Suppose a feather appears green but has no green pigment. Explain in terms of diffraction.
18. It is possible that there is no minimum in the interference pattern of a single slit. Explain why. Is the same true of double slits and diffraction gratings?

27.5 Single Slit Diffraction
19. As the width of the slit producing a single-slit diffraction pattern is reduced, how will the diffraction pattern produced change?

27.6 Limits of Resolution: The Rayleigh Criterion
20. A beam of light always spreads out. Why can a beam not be created with parallel rays to prevent spreading? Why can lenses, mirrors, or apertures not be used to correct the spreading?

27.7 Thin Film Interference
21. What effect does increasing the wedge angle have on the spacing of interference fringes? If the wedge angle is too large, fringes are not observed. Why?
22. How is the difference in paths taken by two originally in-phase light waves related to whether they interfere constructively or destructively? How can this be affected by reflection? By refraction?
23. Is there a phase change in the light reflected from either surface of a contact lens floating on a person's tear layer? The index of refraction of the lens is about 1.5, and its top surface is dry.
24. In placing a sample on a microscope slide, a glass cover is placed over a water drop on the glass slide. Light incident from above can reflect from the top and bottom of the glass cover and from the glass slide below the water drop. At which surfaces will there be a phase change in the reflected light?
25. Answer the above question if the fluid between the two pieces of crown glass is carbon disulfide.
26. While contemplating the food value of a slice of ham, you notice a rainbow of color reflected from its moist surface. Explain its origin.
27. An inventor notices that a soap bubble is dark at its thinnest and realizes that destructive interference is taking place for all wavelengths. How could she use this knowledge to make a non-reflective coating for lenses that is effective at all wavelengths? That is, what limits would there be on the index of refraction and thickness of the coating? How might this be impractical?
28. A non-reflective coating like the one described in Example 27.6 works ideally for a single wavelength and for perpendicular incidence. What happens for other wavelengths and other incident directions? Be specific.
29. Why is it much more difficult to see interference fringes for light reflected from a thick piece of glass than from a thin film? Would it be easier if monochromatic light were used?

27.8 Polarization
30. Under what circumstances is the phase of light changed by reflection? Is the phase related to polarization?
31. Can a sound wave in air be polarized? Explain.
32. No light passes through two perfect polarizing filters with perpendicular axes. However, if a third polarizing filter is placed between the original two, some light can pass. Why is this? Under what circumstances does most of the light pass?
33. Explain what happens to the energy carried by light that it is dimmed by passing it through two crossed polarizing filters.
34. When particles scattering light are much smaller than its wavelength, the amount of scattering is proportional to \(1/\lambda^4\). Does this mean there is more scattering for small \(\lambda\) than large \(\lambda\)? How does this relate to the fact that the sky is blue?
35. Using the information given in the preceding question, explain why sunsets are red.
36. When light is reflected at Brewster's angle from a smooth surface, it is 100% polarized parallel to the surface. Part of the light will be refracted into the surface. Describe how you would do an experiment to determine the polarization of the refracted light. What direction would you expect the polarization to have and would you expect it to be 100%?

27.9 *Extended Topic* Microscopy Enhanced by the Wave Characteristics of Light
37. Explain how microscopes can use wave optics to improve contrast and why this is important.
38. A bright white light under water is collimated and directed upon a prism. What range of colors does one see emerging?
27.1 The Wave Aspect of Light: Interference 
1. Show that when light passes from air to water, its wavelength decreases to 0.75 times its original value.
2. Find the range of visible wavelengths of light in crown glass.
3. What is the index of refraction of a material for which the wavelength of light is 0.671 times its value in a vacuum? Identify the likely substance.
4. Analysis of an interference effect in a clear solid shows that the wavelength of light in the solid is 329 nm. Knowing this light comes from a He-Ne laser and has a wavelength of 633 nm in air, is the substance zircon or diamond?
5. What is the ratio of thicknesses of crown glass and water that would contain the same number of wavelengths of light?

27.3 Young's Double Slit Experiment
6. At what angle is the first-order maximum for 450-nm wavelength blue light falling on double slits separated by 0.0500 mm?
7. Calculate the angle for the third-order maximum of 580-nm wavelength yellow light falling on double slits separated by 0.100 mm.
8. What is the separation between two slits for which 610-nm orange light has its first maximum at an angle of 30.0º?
9. Find the distance between two slits that produces the first minimum for 410-nm violet light at an angle of 45.0º.
10. Calculate the wavelength of light that has its third minimum at an angle of 30.0º when falling on double slits separated by 3.00 µm.
Explicitly, show how you follow the steps in Problem-Solving Strategies for Wave Optics.
11. What is the wavelength of light falling on double slits separated by 2.00 µm if the third-order maximum is at an angle of 60.0º?
12. At what angle is the fourth-order maximum for the situation in Exercise 27.6?
13. What is the highest-order maximum for 400-nm light falling on double slits separated by 25.0 µm?
14. Find the largest wavelength of light falling on double slits separated by 1.20 µm for which there is a first-order maximum. Is this in the visible part of the spectrum?
15. What is the smallest separation between two slits that will produce a second-order maximum for 720-nm red light?
16. (a) What is the smallest separation between two slits that will produce a second-order maximum for any visible light? (b) For all visible light?
17. (a) If the first-order maximum for pure-wavelength light falling on a double slit is at an angle of 10.0º, at what angle is the second-order maximum? (b) What is the angle of the first minimum? (c) What is the highest-order maximum possible here?
18. Figure 27.56 shows a double slit located a distance x from a screen, with the distance from the center of the screen given by y.

When the distance d between the slits is relatively large, there will be numerous bright spots, called fringes. Show that, for small angles (where \( \sin \theta \approx \theta \), with \( \theta \) in radians), the distance between fringes is given by \( \Delta y = \frac{x\lambda}{d} \).

27.4 Multiple Slit Diffraction
21. A diffraction grating has 2000 lines per centimeter. At what angle will the first-order maximum be for 520-nm-wavelength green light?
22. Find the angle for the third-order maximum for 580-nm-wavelength yellow light falling on a diffraction grating having 1500 lines per centimeter.
23. How many lines per centimeter are there on a diffraction grating that gives a first-order maximum for 470-nm blue light at an angle of 25.0º?
24. What is the distance between lines on a diffraction grating that produces a second-order maximum for 760-nm red light at an angle of 60.0º?
25. Calculate the wavelength of light that has its second-order maximum at 45.0º when falling on a diffraction grating that has 5000 lines per centimeter.
26. An electric current through hydrogen gas produces several distinct wavelengths of visible light. What are the wavelengths of the hydrogen spectrum, if they form first-order maxima at angles of 24.2º, 25.7º, 29.1º, and 41.0º when projected on a diffraction grating having 10,000 lines per centimeter? Explicitly show how you follow the steps in Problem-Solving Strategies for Wave Optics.
27. (a) What do the four angles in the above problem become if a 5000-line-per-centimeter diffraction grating is used? (b) Using this grating, what would the angles be for the second-order maxima? (c) Discuss the relationship between integral reductions in lines per centimeter and the new angles of various order maxima.
28. What is the maximum number of lines per centimeter a diffraction grating can have and produce a complete first-order spectrum for visible light?
29. The yellow light from a sodium vapor lamp seems to be of pure wavelength, but it produces two first-order maxima at 36.093º and 36.129º when projected on a 10,000 line per centimeter diffraction grating. What are the two wavelengths to an accuracy of 0.1 nm?
30. What is the spacing between structures in a feather that acts as a reflection grating, given that they produce a first-order maximum for 525-nm light at a 30.0º angle?
31. Structures on a bird feather act like a reflection grating having 8000 lines per centimeter. What is the angle of the first-order maximum for 600-nm light?
32. An opal such as that shown in Figure 27.17 acts like a reflection grating with rows separated by about 8 µm. If the opal is illuminated normally, (a) at what angle will red light be seen and (b) at what angle will blue light be seen?

33. At what angle does a diffraction grating produce a second-order maximum for light having a first-order maximum at 20.0°?

34. Show that a diffraction grating cannot produce a second-order maximum for a given wavelength of light unless the first-order maximum is at an angle less than 30.0°.

35. If a diffraction grating produces a first-order maximum for the shortest wavelength of visible light at 30.0°, at what angle will the first-order maximum be for the longest wavelength of visible light?

36. (a) Find the maximum number of lines per centimeter a diffraction grating can have and produce a maximum for the smallest wavelength of visible light. (b) Would such a grating be useful for ultraviolet spectra? (c) For infrared spectra?

37. (a) Show that a 30,000-line-per-centimeter grating will not produce a maximum for visible light. (b) What is the longest wavelength for which it does produce a first-order maximum? (c) What is the greatest number of lines per centimeter a diffraction grating can have and produce a complete second-order spectrum for visible light?

38. A He–Ne laser beam is reflected from the surface of a CD onto a wall. The brightest spot is the reflected beam at an angle equal to the angle of incidence. However, fringes are also observed. If the wall is 1.50 m from the CD, and the first fringe is 0.600 m from the central maximum, what is the spacing of grooves on the CD?

39. The analysis shown in the figure below also applies to diffraction gratings with lines separated by a distance \( d \). What is the distance between fringes produced by a diffraction grating having 125 lines per centimeter for 600-nm light, if the screen is 1.50 m away?

![Figure 27.57](image)

Figure 27.57 The distance between adjacent fringes is \( \Delta y = x\lambda/d \), assuming the slit separation \( d \) is large compared with \( \lambda \).

40. Unreasonable Results

Red light of wavelength of 700 nm falls on a double slit separated by 400 nm. (a) At what angle is the first minimum in the diffraction pattern? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

41. Unreasonable Results

(a) What visible wavelength has its fourth-order maximum at an angle of 25.0° when projected on a 25,000-line-per-centimeter diffraction grating? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

42. Construct Your Own Problem

Consider a spectrometer based on a diffraction grating. Construct a problem in which you calculate the distance between two wavelengths of electromagnetic radiation in your spectrometer. Among the things to be considered are the wavelengths you wish to be able to distinguish, the number of lines per meter on the diffraction grating, and the distance from the grating to the screen or detector. Discuss the practicality of the device in terms of being able to discern between wavelengths of interest.

27.5 Single slit Diffraction

43. (a) At what angle is the first minimum for 550-nm light falling on a single slit of width 1.00 µm? (b) Will there be a second minimum?

44. (a) Calculate the angle at which a 2.00-µm-wide slit produces its first minimum for 410-nm violet light. (b) Where is the first minimum for 700-nm red light?

45. (a) How wide is a single slit that produces its first minimum for 633-nm light at an angle of 28.0°? (b) At what angle will the second minimum be?

46. (a) What is the width of a single slit that produces its first minimum at 60.0° for 600-nm light? (b) Find the wavelength of light that has its first minimum at 62.0°.

47. Find the wavelength of light that has its third minimum at an angle of 48.6° when it falls on a single slit of width 3.00 µm.

48. Calculate the wavelength of light that produces its first minimum at an angle of 36.9° when falling on a single slit of width 1.00 µm.

49. (a) Sodium vapor light averaging 589 nm in wavelength falls on a single slit of width 7.50 µm. At what angle does it produce its second minimum? (b) What is the highest-order minimum produced?

50. (a) Find the angle of the third diffraction minimum for 633-nm light falling on a slit of width 20.0 µm. (b) What slit width would place this minimum at 85.0°? Explicitly show how you follow the steps in Problem-Solving Strategies for Wave Optics

51. (a) Find the angle between the first minima for the two sodium vapor lines, which have wavelengths of 589.1 and 589.6 nm, when they fall upon a single slit of width 2.00 µm. (b) What is the distance between these minima if the diffraction pattern falls on a screen 1.00 m from the slit? (c) Discuss the ease or difficulty of measuring such a distance.

52. (a) What is the minimum width of a single slit (in multiples of \( \lambda \)) that will produce a first minimum for a wavelength \( \lambda \)? (b) What is its minimum width if it produces 50 minima? (c) 1000 minima?

53. (a) If a single slit produces a first minimum at 14.5°, at what angle is the second-order minimum? (b) What is the angle of the third-order minimum? (c) Is there a fourth-order minimum? (d) Use your answers to illustrate how the angular width of the central maximum is about twice the angular width of the next maximum (which is the angle between the first and second minima).

54. A double slit produces a diffraction pattern that is a combination of single and double slit interference. Find the ratio of the width of the slits to the separation between them, if the first minimum of the single slit pattern falls on the fifth maximum of the double slit pattern. (This will greatly reduce the intensity of the fifth maximum.)

55. Integrated Concepts

A water break at the entrance to a harbor consists of a rock barrier with a 50.0-m-wide opening. Ocean waves of 20.0-m wavelength approach the opening straight on. At what angle to the incident direction are the boats inside the harbor most protected against wave action?

56. Integrated Concepts

An airport maintenance technician walks past a tall hangar door that acts like a single slit for sound entering the hangar. Outside the door, on a line perpendicular to the opening in the door, a jet engine makes a 600-Hz sound. At what angle with the door will the technician observe the first minimum in sound intensity if the vertical opening is 0.800 m wide and the speed of sound is 340 m/s?

27.6 Limits of Resolution: The Rayleigh Criterion
57. The 300-m-diameter Arecibo radio telescope pictured in Figure 27.28 detects radio waves with a 4.00 cm average wavelength.

(a) What is the angle between two just-resolvable point sources for this telescope?

(b) How close together could these point sources be at the 2 million light year distance of the Andromeda galaxy?

58. Assuming the angular resolution found for the Hubble Telescope in Example 27.5, what is the smallest detail that could be observed on the Moon?

59. Diffraction spreading for a flashlight is insignificant compared with other limitations in its optics, such as spherical aberrations in its mirror. To show this, calculate the minimum angular spreading of a flashlight beam that is originally 5.00 cm in diameter with an average wavelength of 600 nm.

60. (a) What is the minimum angular spread of a 633-nm wavelength He-Ne laser beam that is originally 1.00 mm in diameter?

(b) If this laser is aimed at a mountain cliff 15.0 km away, how big will the illuminated spot be?

(c) How big a spot would be illuminated on the Moon, neglecting atmospheric effects? (This might be done to hit a corner reflector to measure the round-trip time and, hence, distance.) Explicitly show how you follow the steps in Problem-Solving Strategies for Wave Optics.

61. A telescope can be used to enlarge the diameter of a laser beam and limit diffraction spreading. The laser beam is sent through the telescope in opposite the normal direction and can then be projected onto a satellite or the Moon.

(a) If this is done with the Mount Wilson telescope, producing a 2.54-m-diameter beam of 633-nm light, what is the minimum angular spread of the beam?

(b) Neglecting atmospheric effects, what is the size of the spot this beam would make on the Moon, assuming a lunar distance of 3.84 × 10^8 m?

62. The limit to the eye's acuity is actually related to diffraction by the pupil.

(a) What is the angle between two just-resolvable points of light for a 3.00-mm-diameter pupil, assuming an average wavelength of 550 nm?

(b) Take your result to be the practical limit for the eye. What is the greatest possible distance a car can be from you if you can resolve its two headlights, given they are 1.30 m apart?

(c) What is the distance between two just-resolvable points held at an arm's length (0.800 m) from your eye?

(d) How does your answer to (c) compare to details you normally observe in everyday circumstances?

63. What is the minimum diameter mirror on a telescope that would allow you to see details as small as 5.00 km on the Moon some 384,000 km away? Assume an average wavelength of 550 nm for the light received.

64. You are told not to shoot until you see the whites of their eyes. If the eyes are separated by 8.5 cm and the diameter of your pupil is 5.0 mm, at what distance can you resolve the two eyes using light of wavelength 555 nm?

65. (a) The planet Pluto and its Moon Charon are separated by 19,600 km. Neglecting atmospheric effects, should the 5.08-m-diameter Mount Palomar telescope be able to resolve these bodies when they are 4.50 × 10^9 km from Earth? Assume an average wavelength of 550 nm.

(b) In actuality, it is just barely possible to discern that Pluto and Charon are separate bodies using an Earth-based telescope. What are the reasons for this?

66. The headlights of a car are 1.3 m apart. What is the maximum distance at which the eye can resolve these two headlights? Take the pupil diameter to be 0.40 cm.

67. When dots are placed on a page from a laser printer, they must be close enough so that you do not see the individual dots of ink. To do this, the separation of the dots must be less than Raleigh's criterion. Take the pupil of the eye to be 3.0 mm and the distance from the paper to the eye of 35 cm; find the minimum separation of two dots such that they cannot be resolved. How many dots per inch (dpi) does this correspond to?

68. Unreasonable Results

An amateur astronomer wants to build a telescope with a diffraction limit that will allow him to see if there are people on the moons of Jupiter.

(a) What diameter mirror is needed to be able to see 1.00 m detail on a Jovian Moon at a distance of 7.50 × 10^8 km from Earth? The wavelength of light averages 600 nm.

(b) What is unreasonable about this result?

(c) Which assumptions are unreasonable or inconsistent?

69. Construct Your Own Problem

Consider diffraction limits for an electromagnetic wave interacting with a circular object. Construct a problem in which you calculate the limit of angular resolution with a device, using this circular object (such as a lens, mirror, or antenna) to make observations. Also calculate the limit to spatial resolution (such as the size of features observable on the Moon) to perform calculations at a specific distance from the device. Among the things to be considered are the wavelength of electromagnetic radiation used, the size of the circular object, and the distance to the system or phenomenon being observed.

27.7 Thin Film Interference

70. A soap bubble is 100 nm thick and illuminated by white light incident perpendicular to its surface. What wavelength and color of visible light is most constructively reflected, assuming the same index of refraction as water?

71. An oil slick on water is 120 nm thick and illuminated by white light incident perpendicular to its surface. What color does the oil appear (what is the most constructively reflected wavelength), given its index of refraction is 1.40?

72. Calculate the minimum thickness of an oil slick on water that appears blue when illuminated by white light perpendicular to its surface. Take the blue wavelength to be 470 nm and the index of refraction of oil to be 1.40.

73. Find the minimum thickness of a soap bubble that appears red when illuminated by white light perpendicular to its surface. Take the wavelength to be 680 nm, and assume the same index of refraction as water.

74. A film of soapy water (n = 1.33) on top of a plastic cutting board has a thickness of 233 nm. What color is most strongly reflected if it is illuminated perpendicular to its surface?

75. What are the three smallest non-zero thicknesses of soapy water (n = 1.33) on Plexiglas if it appears green (constructively reflecting 520-nm light) when illuminated perpendicularly by white light? Explicitly show how you follow the steps in Problem Solving Strategies for Wave Optics.

76. Suppose you have a lens system that is to be used primarily for 700-nm red light. What is the second thinnest coating of fluorite (magnesium fluoride) that would be non-reflective for this wavelength?

77. (a) As a soap bubble thins it becomes dark, because the path length difference becomes small compared with the wavelength of light and there is a phase shift at the top surface. If it becomes dark when the path length difference is less than one-fourth the wavelength, what is the thickest the bubble can be and appear dark at all visible wavelengths? Assume the same index of refraction as water. (b) Discuss the fragility of the film considering the thickness found.

78. A film of oil on water will appear dark when it is very thin, because the path length difference becomes small compared with the wavelength of light and there is a phase shift at the top surface. If it becomes dark when the path length difference is less than one-fourth the wavelength, what is the thickest the oil can be and appear dark at all visible wavelengths? Oil has an index of refraction of 1.40.

79. Figure 27.34 shows two glass slides illuminated by pure-wavelength light incident perpendicularly. The top slide touches the bottom slide at
one end and rests on a 0.100-mm-diameter hair at the other end, forming a wedge of air. (a) How far apart are the dark bands, if the slides are 7.50 cm long and 589-nm light is used? (b) Is there any difference if the slides are made from crown or flint glass? Explain.

80. **Figure 27.34** shows two 7.50-cm-long glass slides illuminated by pure 589-nm wavelength light incident perpendicularly. The top slide touches the bottom slide at one end and rests on some debris at the other end, forming a wedge of air. How thick is the debris, if the dark bands are 1.00 mm apart?

81. Repeat Exercise 27.70, but take the light to be incident at a 45º angle.

82. Repeat Exercise 27.71, but take the light to be incident at a 45º angle.

83. **Unreasonable Results**

To save money on making military aircraft invisible to radar, an inventor decides to coat them with a non-reflective material having an index of refraction of 1.20, which is between that of air and the surface of the plane. This, he reasons, should be much cheaper than designing Stealth bombers. (a) What thickness should the coating be to inhibit the reflection of 4.00-cm wavelength radar? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

27.8 Polarization

84. What angle is needed between the direction of polarized light and the axis of a polarizing filter to cut its intensity in half?

85. The angle between the axes of two polarizing filters is 45.0º. By how much does the second filter reduce the intensity of the light coming through the first?

86. If you have completely polarized light of intensity 150 W/m², what will its intensity be after passing through a polarizing filter with its axis at an angle of 89.0º to the light's polarization direction?

87. What angle would the axis of a polarizing filter need to make with the direction of polarized light of intensity 1.00 kW/m² to reduce the intensity to 10.0 W/m²?

88. At the end of Example 27.8, it was stated that the intensity of polarized light is reduced to 90.0% of its original value by passing through a polarizing filter with its axis at an angle of 18.4º to the direction of polarization. Verify this statement.

89. Show that if you have three polarizing filters, with the second at an angle of 45º to the first and the third at an angle of 90.0º to the first, the intensity of light passed by the first will be reduced to 25.0% of its value. (This is in contrast to having only the first and third, which reduces the intensity to zero, so that placing the second between them increases the intensity of the transmitted light.)

90. Prove that, if I is the intensity of light transmitted by two polarizing filters with axes at an angle 0 and I’ is the intensity when the axes are at an angle 90.0º–0, then I + I’ = I₀, the original intensity. (Hint: Use the trigonometric identities cos(90.0º–0) = sin 0 and

cos²0 + sin²0 = 1.)

91. At what angle will light reflected from diamond be completely polarized?

92. What is Brewster's angle for light traveling in water that is reflected from crown glass?

93. A scuba diver sees light reflected from the water's surface. At what angle will this light be completely polarized?

94. At what angle is light inside crown glass completely polarized when reflected from water, as in a fish tank?
Figure 28.1 Special relativity explains why traveling to other star systems, such as these in the Orion Nebula, is unreasonable using our current level of technology. (credit: s58y, Flickr)

<table>
<thead>
<tr>
<th>Learning Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>28.1. Einstein’s Postulates</strong></td>
</tr>
<tr>
<td>• State and explain both of Einstein’s postulates.</td>
</tr>
<tr>
<td>• Explain what an inertial frame of reference is.</td>
</tr>
<tr>
<td>• Describe one way the speed of light can be changed.</td>
</tr>
<tr>
<td><strong>28.2. Simultaneity And Time Dilation</strong></td>
</tr>
<tr>
<td>• Describe simultaneity.</td>
</tr>
<tr>
<td>• Describe time dilation.</td>
</tr>
<tr>
<td>• Calculate γ.</td>
</tr>
<tr>
<td>• Compare proper time and the observer’s measured time.</td>
</tr>
<tr>
<td>• Explain why the twin paradox is a false paradox.</td>
</tr>
<tr>
<td><strong>28.3. Length Contraction</strong></td>
</tr>
<tr>
<td>• Describe proper length.</td>
</tr>
<tr>
<td>• Calculate length contraction.</td>
</tr>
<tr>
<td>• Explain why we don’t notice these effects at everyday scales.</td>
</tr>
<tr>
<td><strong>28.4. Relativistic Addition of Velocities</strong></td>
</tr>
<tr>
<td>• Calculate relativistic velocity addition.</td>
</tr>
<tr>
<td>• Explain when relativistic velocity addition should be used instead of classical addition of velocities.</td>
</tr>
<tr>
<td>• Calculate relativistic Doppler shift.</td>
</tr>
<tr>
<td><strong>28.5. Relativistic Momentum</strong></td>
</tr>
<tr>
<td>• Calculate relativistic momentum.</td>
</tr>
<tr>
<td>• Explain why the only mass it makes sense to talk about is rest mass.</td>
</tr>
<tr>
<td><strong>28.6. Relativistic Energy</strong></td>
</tr>
<tr>
<td>• Compute total energy of a relativistic object.</td>
</tr>
<tr>
<td>• Compute the kinetic energy of a relativistic object.</td>
</tr>
<tr>
<td>• Describe rest energy, and explain how it can be converted to other forms.</td>
</tr>
<tr>
<td>• Explain why massive particles cannot reach C.</td>
</tr>
</tbody>
</table>

**Introduction to Special Relativity**

Have you ever looked up at the night sky and dreamed of traveling to other planets in faraway star systems? Would there be other life forms? What would other worlds look like? You might imagine that such an amazing trip would be possible if we could just travel fast enough, but you will read in this chapter why this is not true. In 1905 Albert Einstein developed the theory of special relativity. This theory explains the limit on an object’s speed and describes the consequences.

Relativity. The word *relativity* might conjure an image of Einstein, but the idea did not begin with him. People have been exploring relativity for many centuries. Relativity is the study of how different observers measure the same event. Galileo and Newton developed the first correct version of classical relativity. Einstein developed the modern theory of relativity. Modern relativity is divided into two parts. *Special relativity* deals with observers who are moving at constant velocity. *General relativity* deals with observers who are undergoing acceleration. Einstein is famous because his theories of relativity made revolutionary predictions. Most importantly, his theories have been verified to great precision in a vast range of experiments, altering forever our concept of space and time.
It is important to note that although classical mechanics, in general, and classical relativity, in particular, are limited, they are extremely good approximations for large, slow-moving objects. Otherwise, we could not use classical physics to launch satellites or build bridges. In the classical limit (objects larger than submicroscopic and moving slower than about 1% of the speed of light), relativistic mechanics becomes the same as classical mechanics. This fact will be noted at appropriate places throughout this chapter.

28.1 Einstein's Postulates

Einstein essentially did the theoretical aspect of this method for relativity. With two deceptively simple postulates and a careful consideration of how measurements are made, he produced the theory of special relativity.

Einstein's First Postulate

The first postulate upon which Einstein based the theory of special relativity relates to reference frames. All velocities are measured relative to some frame of reference. For example, a car's motion is measured relative to its starting point or the road it is moving over, a projectile's motion is measured relative to the surface it was launched from, and a planet's orbit is measured relative to the star it is orbiting around. The simplest frames of reference are those that are not accelerated and are not rotating. Newton's first law, the law of inertia, holds exactly in such a frame.

An inertial frame of reference is a reference frame in which a body at rest remains at rest and a body in motion moves at a constant speed in a straight line unless acted on by an outside force.

The laws of physics seem to be simplest in inertial frames. For example, when you are in a plane flying at a constant altitude and speed, physics seems to work exactly the same as if you were standing on the surface of the Earth. However, in a plane that is taking off, matters are somewhat more complicated. In these cases, the net force on an object, \( F \), is not equal to the product of mass and acceleration, \( ma \). Instead, \( F \) is equal to \( ma \) plus a fictitious force. This situation is not as simple as in an inertial frame. Not only are laws of physics simplest in inertial frames, but they
should be the same in all inertial frames, since there is no preferred frame and no absolute motion. Einstein incorporated these ideas into his first postulate of special relativity.

**First Postulate of Special Relativity**

The laws of physics are the same and can be stated in their simplest form in all inertial frames of reference.

As with many fundamental statements, there is more to this postulate than meets the eye. The laws of physics include only those that satisfy this postulate. We shall find that the definitions of relativistic momentum and energy must be altered to fit. Another outcome of this postulate is the famous equation \( E = mc^2 \).

**Einstein's Second Postulate**

The second postulate upon which Einstein based his theory of special relativity deals with the speed of light. Late in the 19th century, the major tenets of classical physics were well established. Two of the most important were the laws of electricity and magnetism and Newton's laws. In particular, the laws of electricity and magnetism predict that light travels at \( c = 3.00 \times 10^8 \text{ m/s} \) in a vacuum, but they do not specify the frame of reference in which light has this speed.

There was a contradiction between this prediction and Newton's laws, in which velocities add like simple vectors. If the latter were true, then two observers moving at different speeds would see light traveling at different speeds. Imagine what a light wave would look like to a person traveling along with it at a speed \( c \). If such a motion were possible then the wave would be stationary relative to the observer. It would have electric and magnetic fields that varied in strength at various distances from the observer but were constant in time. This is not allowed by Maxwell's equations. So either Maxwell's equations are wrong, or an object with mass cannot travel at speed \( c \). Einstein concluded that the latter is true. An object with mass cannot travel at speed \( c \), and the further implication is that light in a vacuum must always travel at speed \( c \) relative to any observer. Maxwell's equations are correct, and Newton's addition of velocities is not correct for light.

Investigations such as Young's double slit experiment in the early-1800s had convincingly demonstrated that light is a wave. Many types of waves were known, and all travelled in some medium. Scientists therefore assumed that a medium carried light, even in a vacuum, and light travelled at a speed \( c \) relative to that medium. Starting in the mid-1880s, the American physicist A. A. Michelson, later aided by E. W. Morley, made a series of direct measurements of the speed of light. The results of their measurements were startling.

**Michelson-Morley Experiment**

The Michelson-Morley experiment demonstrated that the speed of light in a vacuum is independent of the motion of the Earth about the Sun.

The eventual conclusion derived from this result is that light, unlike mechanical waves such as sound, does not need a medium to carry it. Furthermore, the Michelson-Morley results implied that the speed of light \( c \) is independent of the motion of the source relative to the observer. That is, everyone observes light to move at speed \( c \) regardless of how they move relative to the source or one another. For a number of years, many scientists tried unsuccessfully to explain these results and still retain the general applicability of Newton's laws.

It was not until 1905, when Einstein published his first paper on special relativity, that the currently accepted conclusion was reached. Based mostly on his analysis that the laws of electricity and magnetism would not allow another speed for light, and only slightly aware of the Michelson-Morley experiment, Einstein detailed his second postulate of special relativity.

**Second Postulate of Special Relativity**

The speed of light \( c \) is a constant, independent of the relative motion of the source and observer.

Deceptively simple and counterintuitive, this and the first postulate leave all else open for change. Some fundamental concepts do change. Among the changes are the loss of agreement on the elapsed time for an event, the variation of distance with speed, and the realization that matter and energy can be converted into one another. You will read about these concepts in the following sections.

**Misconception Alert: Constancy of the Speed of Light**

The speed of light is a constant \( c = 3.00 \times 10^8 \text{ m/s} \) in a vacuum. If you remember the effect of the index of refraction from The Law of Refraction, the speed of light is lower in matter.

**Check Your Understanding**

Explain how special relativity differs from general relativity.

**Solution**

Special relativity applies only to unaccelerated motion, but general relativity applies to accelerated motion.
28.2 Simultaneity And Time Dilation

Do time intervals depend on who observes them? Intuitively, we expect the time for a process, such as the elapsed time for a foot race, to be the same for all observers. Our experience has been that disagreements over elapsed time have to do with the accuracy of measuring time. When we carefully consider just how time is measured, however, we will find that elapsed time depends on the relative motion of an observer with respect to the process being measured.

Simultaneity

Consider how we measure elapsed time. If we use a stopwatch, for example, how do we know when to start and stop the watch? One method is to use the arrival of light from the event, such as observing a light turning green to start a drag race. The timing will be more accurate if some sort of electronic detection is used, avoiding human reaction times and other complications.

Now suppose we use this method to measure the time interval between two flashes of light produced by flash lamps. (See Figure 28.5.) Two flash lamps with observer A midway between them are on a rail car that moves to the right relative to observer B. The light flashes are emitted just as A passes B, so that both A and B are equidistant from the lamps when the light is emitted. Observer B measures the time interval between the arrival of the light flashes. According to postulate 2, the speed of light is not affected by the motion of the lamps relative to B. Therefore, light travels equal distances to him at equal speeds. Thus observer B measures the flashes to be simultaneous.

Simultaneity is not absolute.

Now consider what observer B sees happen to observer A. She receives the light from the right first, because she has moved towards that flash lamp, lessening the distance the light must travel and reducing the time it takes to get to her. Light travels at speed \( c \) relative to both observers, but observer B remains equidistant from the points where the flashes were emitted, while A gets closer to the emission point on the right. From observer B’s point of view, then, there is a time interval between the arrival of the flashes to observer A. Observer B measures the flashes to be simultaneous relative to him but not relative to A. Here a relative velocity between observers affects whether two events are observed to be simultaneous. Simultaneity is not absolute.

This illustrates the power of clear thinking. We might have guessed incorrectly that if light is emitted simultaneously, then two observers halfway between the sources would see the flashes simultaneously. But careful analysis shows this not to be the case. Einstein was brilliant at this type of
thought experiment (in German, “Gedankenexperiment”). He very carefully considered how an observation is made and disregarded what might seem obvious. The validity of thought experiments, of course, is determined by actual observation. The genius of Einstein is evidenced by the fact that experiments have repeatedly confirmed his theory of relativity.

In summary: Two events are defined to be simultaneous if an observer measures them as occurring at the same time (such as by receiving light from the events). Two events are not necessarily simultaneous to all observers.

**Time Dilation**

The consideration of the measurement of elapsed time and simultaneity leads to an important relativistic effect.

**Time dilation**

Time dilation is the phenomenon of time passing slower for an observer who is moving relative to another observer.

Suppose, for example, an astronaut measures the time it takes for light to cross her ship, bounce off a mirror, and return. (See Figure 28.6.) How does the elapsed time the astronaut measures compare with the elapsed time measured for the same event by a person on the Earth? Asking this question (another thought experiment) produces a profound result. We find that the elapsed time for a process depends on who is measuring it. In this case, the time measured by the astronaut is smaller than the time measured by the Earth-bound observer. The passage of time is different for the observers because the distance the light travels in the astronaut’s frame is smaller than in the Earth-bound frame. Light travels at the same speed in each frame, and so it will take longer to travel the greater distance in the Earth-bound frame.

![Diagram of astronaut and Earth-bound observer measuring light travel time](image)

**Figure 28.6** (a) An astronaut measures the time \( \Delta t_0 \) for light to cross her ship using an electronic timer. Light travels a distance \( 2D \) in the astronaut’s frame. (b) A person on the Earth sees the light follow the longer path \( 2s \) and take a longer time \( \Delta t \). (c) These triangles are used to find the relationship between the two distances \( 2D \) and \( 2s \).

To quantitatively verify that time depends on the observer, consider the paths followed by light as seen by each observer. (See Figure 28.6(c).) The astronaut sees the light travel straight across and back for a total distance of \( 2D \), twice the width of her ship. The Earth-bound observer sees the light travel a total distance \( 2s \). Since the ship is moving at speed \( v \) to the right relative to the Earth, light moving to the right hits the mirror in this frame. Light travels at a speed \( c \) in both frames, and because time is the distance divided by speed, the time measured by the astronaut is

\[
\Delta t_0 = \frac{2D}{c}. \tag{28.1}
\]

This time has a separate name to distinguish it from the time measured by the Earth-bound observer.

**Proper Time**

Proper time \( \Delta t_0 \) is the time measured by an observer at rest relative to the event being observed.

In the case of the astronaut observe the reflecting light, the astronaut measures proper time. The time measured by the Earth-bound observer is

\[
\Delta t = \frac{2s}{c}. \tag{28.2}
\]
To find the relationship between \( \Delta t_0 \) and \( \Delta t \), consider the triangles formed by \( D \) and \( s \). (See Figure 28.6(c).) The third side of these similar triangles is \( L \), the distance the astronaut moves as the light goes across her ship. In the frame of the Earth-bound observer,

\[
L = \frac{\sqrt{2}L}{2}.
\]  
(28.3)

Using the Pythagorean Theorem, the distance \( s \) is found to be

\[
s = \sqrt{D^2 + \left(\frac{\sqrt{2}L}{2}\right)^2}.
\]  
(28.4)

Substituting \( s \) into the expression for the time interval \( \Delta t \) gives

\[
\Delta t = \frac{s}{c} = \frac{\sqrt{D^2 + \left(\frac{\sqrt{2}t}{2}\right)^2}}{c}.
\]  
(28.5)

We square this equation, which yields

\[
(\Delta t)^2 = \left(\frac{s}{c}\right)^2 = \frac{4(D^2 + \left(\frac{\sqrt{2}t}{2}\right)^2)}{c^2} = \frac{4D^2}{c^2} + \frac{\sqrt{2}t}{c^2}(\Delta t)^2.
\]  
(28.6)

Note that if we square the first expression we had for \( \Delta t_0 \), we get \( (\Delta t_0)^2 = \frac{4D^2}{c^2} \). This term appears in the preceding equation, giving us a means to relate the two time intervals. Thus,

\[
(\Delta t)^2 = (\Delta t_0)^2 + \frac{\sqrt{2}t}{c^2}(\Delta t)^2.
\]  
(28.7)

Gathering terms, we solve for \( \Delta t \):

\[
(\Delta t)^2\left(1 - \frac{\sqrt{2}}{c^2}\right) = (\Delta t_0)^2.
\]  
(28.8)

Thus,

\[
(\Delta t)^2 = \frac{(\Delta t_0)^2}{1 - \frac{\sqrt{2}}{c^2}}.
\]  
(28.9)

Taking the square root yields an important relationship between elapsed times:

\[
\Delta t = \frac{\Delta t_0}{\sqrt{1 - \frac{\sqrt{2}}{c^2}}} = \gamma \Delta t_0,
\]  
(28.10)

where

\[
\gamma = \frac{1}{\sqrt{1 - \frac{\sqrt{2}}{c^2}}}
\]  
(28.11)

This equation for \( \Delta t \) is truly remarkable. First, as contended, elapsed time is not the same for different observers moving relative to one another, even though both are in inertial frames. Proper time \( \Delta t_0 \) measured by an observer, like the astronaut moving with the apparatus, is smaller than time measured by other observers. Since those other observers measure a longer time \( \Delta t \), the effect is called time dilation. The Earth-bound observer sees time dilate (get longer) for a system moving relative to the Earth. Alternatively, according to the Earth-bound observer, time slows in the moving frame, since less time passes there. All clocks moving relative to an observer, including biological clocks such as aging, are observed to run slow compared with a clock stationary relative to the observer.

Note that if the relative velocity is much less than the speed of light (\( v << c \)), then \( \frac{\sqrt{2}}{c^2} \) is extremely small, and the elapsed times \( \Delta t \) and \( \Delta t_0 \) are nearly equal. At low velocities, modern relativity approaches classical physics—our everyday experiences have very small relativistic effects.

The equation \( \Delta t = \gamma \Delta t_0 \) also implies that relative velocity cannot exceed the speed of light. As \( v \) approaches \( c \), \( \Delta t \) approaches infinity. This would imply that time in the astronaut’s frame stops at the speed of light. If \( v \) exceeded \( c \), then we would be taking the square root of a negative number, producing an imaginary value for \( \Delta t \).

There is considerable experimental evidence that the equation \( \Delta t = \gamma \Delta t_0 \) is correct. One example is found in cosmic ray particles that continuously rain down on the Earth from deep space. Some collisions of these particles with nuclei in the upper atmosphere result in short-lived particles called muons. The half-life (amount of time for half of a material to decay) of a muon is 1.52 \( \mu \)s when it is at rest relative to the observer who measures the
half-life. This is the proper time $\Delta t_0$. Muons produced by cosmic ray particles have a range of velocities, with some moving near the speed of light. It has been found that the muon’s half-life as measured by an Earth-bound observer ($\Delta t$) varies with velocity exactly as predicted by the equation $\Delta t = \gamma \Delta t_0$. The faster the muon moves, the longer it lives. We on the Earth see the muon’s half-life time dilated—as viewed from our frame, the muon decays more slowly than it does when at rest relative to us.

**Example 28.1 Calculating $\Delta t$ for a Relativistic Event: How Long Does a Speedy Muon Live?**

Suppose a cosmic ray colliding with a nucleus in the Earth’s upper atmosphere produces a muon that has a velocity $v = 0.950c$. The muon then travels at constant velocity and lives $1.52\ \mu s$ as measured in the muon’s frame of reference. (You can imagine this as the muon’s internal clock.) How long does the muon live as measured by an Earth-bound observer? (See Figure 28.7.)

**Strategy**

A clock moving with the system being measured observes the proper time, so the time we are given is $\Delta t_0 = 1.52\ \mu s$. The Earth-bound observer measures $\Delta t$ as given by the equation $\Delta t = \gamma \Delta t_0$. Since we know the velocity, the calculation is straightforward.

**Solution**

1) Identify the knowns. $v = 0.950c$, $\Delta t_0 = 1.52\ \mu s$

2) Identify the unknown. $\Delta t$

3) Choose the appropriate equation.

Use,

$$\Delta t = \gamma \Delta t_0, \quad (28.12)$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (28.13)$$

4) Plug the knowns into the equation.

First find $\gamma$.

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$= \frac{1}{\sqrt{1 - \frac{(0.950c)^2}{c^2}}}$$

$$= \frac{1}{\sqrt{1 - (0.950)^2}}$$

$$= 3.20.$$
\[ \Delta t = \gamma \Delta t_0 \]  
\[ = (3.20)(1.52 \ \mu s) \]  
\[ = 4.87 \ \mu s \]  

**Discussion**

One implication of this example is that since \( \gamma = 3.20 \) at 95.0\% of the speed of light (\( v = 0.950c \)), the relativistic effects are significant. The two time intervals differ by this factor of 3.20, where classically they would be the same. Something moving at 0.950c is said to be highly relativistic.

Another implication of the preceding example is that everything an astronaut does when moving at 95.0\% of the speed of light relative to the Earth takes 3.20 times longer when observed from the Earth. Does the astronaut sense this? Only if she looks outside her spaceship. All methods of measuring time in her frame will be affected by the same factor of 3.20. This includes her wristwatch, heart rate, cell metabolism rate, nerve impulse rate, and so on. She will have no way of telling, since all of her clocks will agree with one another because their relative velocities are zero. Motion is relative, not absolute. But what if she does look out the window?

**Real-World Connections**

It may seem that special relativity has little effect on your life, but it is probably more important than you realize. One of the most common effects is through the Global Positioning System (GPS). Emergency vehicles, package delivery services, electronic maps, and communications devices are just a few of the common uses of GPS, and the GPS system could not work without taking into account relativistic effects. GPS satellites rely on precise time measurements to communicate. The signals travel at relativistic speeds. Without corrections for time dilation, the satellites could not communicate, and the GPS system would fail within minutes.

**The Twin Paradox**

An intriguing consequence of time dilation is that a space traveler moving at a high velocity relative to the Earth would age less than her Earth-bound twin. Imagine the astronaut moving at such a velocity that \( \gamma = 30.0 \), as in Figure 28.8. A trip that takes 2.00 years in her frame would take 60.0 years in her Earth-bound twin’s frame. Suppose the astronaut traveled 1.00 year to another star system. She briefly explored the area, and then traveled 1.00 year back. If the astronaut was 40 years old when she left, she would be 42 upon her return. Everything on the Earth, however, would have aged 60.0 years. Her twin, if still alive, would be 100 years old.

The situation would seem different to the astronaut. Because motion is relative, the spaceship would seem to be stationary and the Earth would appear to move. (This is the sensation you have when flying in a jet.) If the astronaut looks out the window of the spaceship, she will see time slow down on the Earth by a factor of \( \gamma = 30.0 \). To her, the Earth-bound sister will have aged only 2/30 (1/15) of a year, while she aged 2.00 years. The two sisters cannot both be correct.

Figure 28.8 The twin paradox asks why the traveling twin ages less than the Earth-bound twin. That is the prediction we obtain if we consider the Earth-bound twin’s frame. In the astronaut’s frame, however, the Earth is moving and time runs slower there. Who is correct?

As with all paradoxes, the premise is faulty and leads to contradictory conclusions. In fact, the astronaut’s motion is significantly different from that of the Earth-bound twin. The astronaut accelerates to a high velocity and then decelerates to view the star system. To return to the Earth, she again accelerates and decelerates. The Earth-bound twin does not experience these accelerations. So the situation is not symmetric, and it is not correct to claim that the astronaut will observe the same effects as her Earth-bound twin. If you use special relativity to examine the twin paradox, you must keep in mind that the theory is expressly based on inertial frames, which by definition are not accelerated or rotating. Einstein developed general relativity to deal with accelerated frames and with gravity, a prime source of acceleration. You can also use general relativity to address the twin paradox and, according to general relativity, the astronaut will age less. Some important conceptual aspects of general relativity are discussed in General Relativity and Quantum Gravity of this course.

In 1971, American physicists Joseph Hafele and Richard Keating verified time dilation at low relative velocities by flying extremely accurate atomic clocks around the Earth on commercial aircraft. They measured elapsed time to an accuracy of a few nanoseconds and compared it with the time...
measured by clocks left behind. Hafele and Keating’s results were within experimental uncertainties of the predictions of relativity. Both special and general relativity had to be taken into account, since gravity and accelerations were involved as well as relative motion.

### Check Your Understanding

1. What is \( \gamma \) if \( v = 0.650c \)?

**Solution**

\[
\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \left(\frac{0.650c}{c}\right)^2}} = 1.32
\]

2. A particle travels at \( 1.90 \times 10^8 \text{ m/s} \) and lives \( 2.10 \times 10^{-8} \text{ s} \) when at rest relative to an observer. How long does the particle live as viewed in the laboratory?

**Solution**

\[
\Delta t = \frac{\Delta t_0}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{2.10 \times 10^{-8} \text{ s}}{\sqrt{1 - \left(\frac{1.90 \times 10^8 \text{ m/s}}{3.00 \times 10^8 \text{ m/s}}\right)^2}} = 2.71 \times 10^{-8} \text{ s}
\]

### 28.3 Length Contraction

**Figure 28.9** People might describe distances differently, but at relativistic speeds, the distances really are different. (credit: Corey Leopold, Flickr)

Have you ever driven on a road that seems like it goes on forever? If you look ahead, you might say you have about 10 km left to go. Another traveler might say the road ahead looks like it’s about 15 km long. If you both measured the road, however, you would agree. Traveling at everyday speeds, the distance you both measure would be the same. You will read in this section, however, that this is not true at relativistic speeds. Close to the speed of light, distances measured are not the same when measured by different observers.

**Proper Length**

One thing all observers agree upon is relative speed. Even though clocks measure different elapsed times for the same process, they still agree that relative speed, which is distance divided by elapsed time, is the same. This implies that distance, too, depends on the observer’s relative motion. If two observers see different times, then they must also see different distances for relative speed to be the same to each of them.

The muon discussed in Example 28.1 illustrates this concept. To an observer on the Earth, the muon travels at \( 0.950c \) for \( 7.05 \mu s \) from the time it is produced until it decays. Thus it travels a distance

\[
L_0 = v\Delta t = (0.950)(3.00\times10^8 \text{ m/s})(7.05\times10^{-6} \text{ s}) = 2.01 \text{ km}
\]

(28.16)

relative to the Earth. In the muon’s frame of reference, its lifetime is only \( 2.20 \mu s \). It has enough time to travel only

\[
L = v\Delta t_0 = (0.950)(3.00\times10^8 \text{ m/s})(2.20\times10^{-6} \text{ s}) = 0.627 \text{ km}.
\]

(28.17)

The distance between the same two events (production and decay of a muon) depends on who measures it and how they are moving relative to it.

**Proper Length**

**Proper length** \( L_0 \) is the distance between two points measured by an observer who is at rest relative to both of the points.

The Earth-bound observer measures the proper length \( L_0 \), because the points at which the muon is produced and decays are stationary relative to the Earth. To the muon, the Earth, air, and clouds are moving, and so the distance \( L \) it sees is not the proper length.
Figure 28.10 (a) The Earth-bound observer sees the muon travel 2.01 km between clouds. (b) The muon sees itself travel the same path, but only a distance of 0.627 km. The Earth, air, and clouds are moving relative to the muon in its frame, and all appear to have smaller lengths along the direction of travel.

Length Contraction

To develop an equation relating distances measured by different observers, we note that the velocity relative to the Earth-bound observer in our muon example is given by

\[ v = \frac{L_0}{\Delta t}. \]  

(28.18)

The time relative to the Earth-bound observer is \( \Delta t \), since the object being timed is moving relative to this observer. The velocity relative to the moving observer is given by

\[ v = \frac{L}{\Delta t_0}. \]  

(28.19)

The moving observer travels with the muon and therefore observes the proper time \( \Delta t_0 \). The two velocities are identical; thus,

\[ \frac{L_0}{\Delta t} = \frac{L}{\Delta t_0}. \]  

(28.20)

We know that \( \Delta t = \gamma \Delta t_0 \). Substituting this equation into the relationship above gives

\[ L = L_0 \gamma. \]  

(28.21)

Substituting for \( \gamma \) gives an equation relating the distances measured by different observers.

Length Contraction

Length contraction \( L \) is the shortening of the measured length of an object moving relative to the observer’s frame.

\[ L = L_0 \sqrt{1 - \frac{v^2}{c^2}}. \]  

(28.22)

Example 28.2 Calculating Length Contraction: The Distance between Stars Contracts when You Travel at High Velocity

Suppose an astronaut, such as the twin discussed in Simultaneity and Time Dilation, travels so fast that \( \gamma = 30.00 \). (a) She travels from the Earth to the nearest star system, Alpha Centauri, 4.300 light years (ly) away as measured by an Earth-bound observer. How far apart are the Earth and Alpha Centauri as measured by the astronaut? (b) In terms of \( c \), what is her velocity relative to the Earth? You may neglect the motion of the Earth relative to the Sun. (See Figure 28.11.)
Figure 28.11 (a) The Earth-bound observer measures the proper distance between the Earth and the Alpha Centauri. (b) The astronaut observes a length contraction, since the Earth and the Alpha Centauri move relative to her ship. She can travel this shorter distance in a smaller time (her proper time) without exceeding the speed of light.

**Strategy**

First note that a light year (ly) is a convenient unit of distance on an astronomical scale—it is the distance light travels in a year. For part (a), note that the 4.300 ly distance between the Alpha Centauri and the Earth is the proper distance \( L_0 \), because it is measured by an Earth-bound observer to whom both stars are (approximately) stationary. To the astronaut, the Earth and the Alpha Centauri are moving by at the same velocity, and so the distance between them is the contracted length \( L \). In part (b), we are given \( \gamma \), and so we can find \( v \) by rearranging the definition of \( \gamma \) to express \( v \) in terms of \( c \).

**Solution for (a)**

1. Identify the known. \( L_0 = 4.300 \text{ ly} \); \( \gamma = 30.00 \)
2. Identify the unknown. \( L \)
3. Choose the appropriate equation. \( L = \frac{L_0}{\gamma} \)
4. Rearrange the equation to solve for the unknown.

\[
L = \frac{L_0}{\gamma} = \frac{4.300 \text{ ly}}{30.00} = 0.1433 \text{ ly}
\]

**Solution for (b)**

1. Identify the known. \( \gamma = 30.00 \)
2. Identify the unknown. \( v \) in terms of \( c \)
3. Choose the appropriate equation. \( \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \)
4. Rearrange the equation to solve for the unknown.

\[
\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 30.00 = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}
\]

Squaring both sides of the equation and rearranging terms gives

\[
900.0 = \frac{1}{1 - \frac{v^2}{c^2}}
\]

so that

\[
1 - \frac{v^2}{c^2} = \frac{1}{900.0}
\]
and

\[
\frac{v^2}{c^2} = 1 - \frac{1}{900.0} = 0.99888....
\]  

(28.27)

Taking the square root, we find

\[
\frac{v}{c} = 0.99944,
\]

(28.28)

which is rearranged to produce a value for the velocity

\[
v = 0.9994c.
\]

(28.29)

**Discussion**

First, remember that you should not round off calculations until the final result is obtained, or you could get erroneous results. This is especially true for special relativity calculations, where the differences might only be revealed after several decimal places. The relativistic effect is large here \((\gamma=30.00)\), and we see that \(v\) is approaching (not equaling) the speed of light. Since the distance as measured by the astronaut is so much smaller, the astronaut can travel it in much less time in her frame.

People could be sent very large distances (thousands or even millions of light years) and age only a few years on the way if they traveled at extremely high velocities. But, like emigrants of centuries past, they would leave the Earth they know forever. Even if they returned, thousands to millions of years would have passed on the Earth, obliterating most of what now exists. There is also a more serious practical obstacle to traveling at such velocities; immensely greater energies than classical physics predicts would be needed to achieve such high velocities. This will be discussed in Relativistic Energy.

Why don't we notice length contraction in everyday life? The distance to the grocery shop does not seem to depend on whether we are moving or not. Examining the equation \(L = L_0\sqrt{1 - \frac{v^2}{c^2}}\), we see that at low velocities \(v<<c\) the lengths are nearly equal, the classical expectation. But length contraction is real, if not commonly experienced. For example, a charged particle, like an electron, traveling at relativistic velocity has electric field lines that are compressed along the direction of motion as seen by a stationary observer. (See Figure 28.12.) As the electron passes a detector, such as a coil of wire, its field interacts much more briefly, an effect observed at particle accelerators such as the 3 km long Stanford Linear Accelerator (SLAC). In fact, to an electron traveling down the beam pipe at SLAC, the accelerator and the Earth are all moving by and are length contracted. The relativistic effect is so great than the accelerator is only 0.5 m long to the electron. It is actually easier to get the electron beam down the pipe, since the beam does not have to be as precisely aimed to get down a short pipe as it would down one 3 km long. This, again, is an experimental verification of the Special Theory of Relativity.

**Figure 28.12** The electric field lines of a high-velocity charged particle are compressed along the direction of motion by length contraction. This produces a different signal when the particle goes through a coil, an experimentally verified effect of length contraction.

**Check Your Understanding**

A particle is traveling through the Earth's atmosphere at a speed of 0.750\(c\). To an Earth-bound observer, the distance it travels is 2.50 km. How far does the particle travel in the particle's frame of reference?

**Solution**

\[
L = L_0\sqrt{1 - \frac{v^2}{c^2}} = (2.50 \text{ km})\sqrt{1 - \left(\frac{0.750c}{c}\right)^2} = 1.65 \text{ km}
\]

(28.30)
28.4 Relativistic Addition of Velocities

Figure 28.13 The total velocity of a kayak, like this one on the Deerfield River in Massachusetts, is its velocity relative to the water as well as the water’s velocity relative to the riverbank. (credit: abkfenris, Flickr)

If you’ve ever seen a kayak move down a fast-moving river, you know that remaining in the same place would be hard. The river current pulls the kayak along. Pushing the oars back against the water can move the kayak forward in the water, but that only accounts for part of the velocity. The kayak’s motion is an example of classical addition of velocities. In classical physics, velocities add as vectors. The kayak’s velocity is the vector sum of its velocity relative to the water and the water’s velocity relative to the riverbank.

Classical Velocity Addition

For simplicity, we restrict our consideration of velocity addition to one-dimensional motion. Classically, velocities add like regular numbers in one-dimensional motion. (See Figure 28.14.) Suppose, for example, a girl is riding in a sled at a speed 1.0 m/s relative to an observer. She throws a snowball first forward, then backward at a speed of 1.5 m/s relative to the sled. We denote direction with plus and minus signs in one dimension; in this example, forward is positive. Let \( v \) be the velocity of the sled relative to the Earth, \( u \) the velocity of the snowball relative to the Earth-bound observer, and \( u' \) the velocity of the snowball relative to the sled.

Classical Velocity Addition

\[
\mathbf{u} = \mathbf{v} + \mathbf{u}'
\]

Thus, when the girl throws the snowball forward, \( u = 1.0 \text{ m/s} + 1.5 \text{ m/s} = 2.5 \text{ m/s} \). It makes good intuitive sense that the snowball will head towards the Earth-bound observer faster, because it is thrown forward from a moving vehicle. When the girl throws the snowball backward, \( u = 1.0 \text{ m/s} + (-1.5 \text{ m/s}) = -0.5 \text{ m/s} \). The minus sign means the snowball moves away from the Earth-bound observer.
Relativistic Velocity Addition

The second postulate of relativity (verified by extensive experimental observation) says that classical velocity addition does not apply to light. Imagine a car traveling at night along a straight road, as in Figure 28.15. If classical velocity addition applied to light, then the light from the car’s headlights would approach the observer on the sidewalk at a speed \( u = v + c \). But we know that light will move away from the car at speed \( c \) relative to the driver of the car, and light will move towards the observer on the sidewalk at speed \( c \), too.

Figure 28.15 According to experiment and the second postulate of relativity, light from the car’s headlights moves away from the car at speed \( c \) and towards the observer on the sidewalk at speed \( c \). Classical velocity addition is not valid.

Relativistic Velocity Addition

Either light is an exception, or the classical velocity addition formula only works at low velocities. The latter is the case. The correct formula for one-dimensional relativistic velocity addition is

\[
u = \frac{v + u'}{1 + \frac{vu'}{c^2}},
\]

where \( v \) is the relative velocity between two observers, \( u \) is the velocity of an object relative to one observer, and \( u' \) is the velocity relative to the other observer. (For ease of visualization, we often choose to measure \( u \) in our reference frame, while someone moving at \( v \) relative to us measures \( u' \).) Note that the term \( \frac{vu'}{c^2} \) becomes very small at low velocities, and \( u = \frac{v + u'}{1 + \frac{vu'}{c^2}} \) gives a result very close to classical velocity addition. As before, we see that classical velocity addition is an excellent approximation to the correct relativistic formula for small velocities. No wonder that it seems correct in our experience.

Example 28.3 Showing that the Speed of Light towards an Observer is Constant (in a Vacuum): The Speed of Light is the Speed of Light

Suppose a spaceship heading directly towards the Earth at half the speed of light sends a signal to us on a laser-produced beam of light. Given that the light leaves the ship at speed \( c \) as observed from the ship, calculate the speed at which it approaches the Earth.

Figure 28.16

Strategy

Because the light and the spaceship are moving at relativistic speeds, we cannot use simple velocity addition. Instead, we can determine the speed at which the light approaches the Earth using relativistic velocity addition.
Solution
1. Identify the knowns. \( v = 0.500c \); \( u' = c \)
2. Identify the unknown. \( u \)
3. Choose the appropriate equation. \( u = \frac{v + u'}{1 + \frac{vu'}{c^2}} \)
4. Plug the knowns into the equation.

\[
\begin{align*}
u &= \frac{v + u'}{1 + \frac{vu'}{c^2}} \\
&= \frac{0.500c + c}{1 + \frac{(0.500c)(c)}{c^2}} \\
&= \frac{(0.500 + 1)c}{1 + \frac{0.500c^2}{c^2}} \\
&= \frac{1.500c}{1 + 0.500} \\
&= \frac{1.500c}{1.500} \\
&= c
\end{align*}
\]

Discussion
Relativistic velocity addition gives the correct result. Light leaves the ship at speed \( c \) and approaches the Earth at speed \( c \). The speed of light is independent of the relative motion of source and observer, whether the observer is on the ship or Earth-bound.

Velocities cannot add to greater than the speed of light, provided that \( v \) is less than \( c \) and \( u' \) does not exceed \( c \). The following example illustrates that relativistic velocity addition is not as symmetric as classical velocity addition.

Example 28.4 Comparing the Speed of Light towards and away from an Observer: Relativistic Package Delivery

Suppose the spaceship in the previous example is approaching the Earth at half the speed of light and shoots a canister at a speed of \( 0.750c \).

(a) At what velocity will an Earth-bound observer see the canister if it is shot directly towards the Earth? (b) If it is shot directly away from the Earth? (See Figure 28.17.)

Figure 28.17

Strategy
Because the canister and the spaceship are moving at relativistic speeds, we must determine the speed of the canister by an Earth-bound observer using relativistic velocity addition instead of simple velocity addition.

Solution for (a)
1. Identify the knowns. \( v = 0.500c \); \( u' = 0.750c \)
2. Identify the unknown. \( u \)
3. Choose the appropriate equation. \( u = \frac{v + u'}{1 + \frac{vu'}{c^2}} \)
4. Plug the knowns into the equation.
\[ u = \frac{v + u'}{1 + \frac{v u'}{c^2}} \]  
\[ = \frac{0.500c + 0.750c}{1 + \frac{(0.500c)(0.750c)}{c^2}} \]  
\[ = \frac{1.250c}{1 + 0.375} \]  
\[ = 0.909c \]

**Solution for (b)**

1. Identify the knowns. \( v = 0.500c \); \( u' = -0.750c \)
2. Identify the unknown. \( u \)
3. Choose the appropriate equation. \( u = \frac{v + u'}{1 + \frac{v u'}{c^2}} \)
4. Plug the knowns into the equation.

\[ u = \frac{v + u'}{1 + \frac{v u'}{c^2}} \]
\[ = \frac{0.500c + (-0.750c)}{1 + \frac{(0.500c)(-0.750c)}{c^2}} \]  
\[ = -0.250c \]  
\[ = -0.400c \]

**Discussion**

The minus sign indicates velocity away from the Earth (in the opposite direction from \( v \)), which means the canister is heading towards the Earth in part (a) and away in part (b), as expected. But relativistic velocities do not add as simply as they do classically. In part (a), the canister does approach the Earth faster, but not at the simple sum of \( 1.250c \). The total velocity is less than you would get classically. And in part (b), the canister moves away from the Earth at a velocity of \( -0.400c \), which is faster than the \( -0.250c \) you would expect classically. The velocities are not even symmetric. In part (a) the canister moves \( 0.409c \) faster than the ship relative to the Earth, whereas in part (b) it moves \( 0.900c \) slower than the ship.

**Doppler Shift**

Although the speed of light does not change with relative velocity, the frequencies and wavelengths of light do. First discussed for sound waves, a Doppler shift occurs in any wave when there is relative motion between source and observer.

**Relativistic Doppler Effects**

The observed wavelength of electromagnetic radiation is longer (called a red shift) than that emitted by the source when the source moves away from the observer and shorter (called a blue shift) when the source moves towards the observer.

\[ \lambda_{\text{obs}} = \lambda_s \left| \frac{1 + \frac{u}{c}}{1 - \frac{u}{c}} \right| \]  

In the Doppler equation, \( \lambda_{\text{obs}} \) is the observed wavelength, \( \lambda_s \) is the source wavelength, and \( u \) is the relative velocity of the source to the observer. The velocity \( u \) is positive for motion away from an observer and negative for motion toward an observer. In terms of source frequency and observed frequency, this equation can be written

\[ f_{\text{obs}} = f_s \left| \frac{1 - \frac{u}{c}}{1 + \frac{u}{c}} \right| \]

Notice that the – and + signs are different than in the wavelength equation.

**Career Connection: Astronomer**

If you are interested in a career that requires a knowledge of special relativity, there’s probably no better connection than astronomy. Astronomers must take into account relativistic effects when they calculate distances, times, and speeds of black holes, galaxies, quasars, and all other astronomical objects. To have a career in astronomy, you need at least an undergraduate degree in either physics or astronomy, but a Master’s or doctoral degree is often required. You also need a good background in high-level mathematics.
Example 28.5 Calculating a Doppler Shift: Radio Waves from a Receding Galaxy

Suppose a galaxy is moving away from the Earth at a speed $0.825c$. It emits radio waves with a wavelength of $0.525\text{ m}$. What wavelength would we detect on the Earth?

**Strategy**

Because the galaxy is moving at a relativistic speed, we must determine the Doppler shift of the radio waves using the relativistic Doppler shift instead of the classical Doppler shift.

**Solution**

1. Identify the knowns. $u=0.825c$; $\lambda_s = 0.525\text{ m}$
2. Identify the unknown. $\lambda_{\text{obs}}$
3. Choose the appropriate equation. $\lambda_{\text{obs}} = \lambda_s \sqrt{\frac{1+\frac{u}{c}}{1-\frac{u}{c}}}$
4. Plug the knowns into the equation.

$$\lambda_{\text{obs}} = \lambda_s \sqrt{\frac{1+\frac{u}{c}}{1-\frac{u}{c}}} = \left(0.525\text{ m}\right) \sqrt{\frac{1+\frac{0.825c}{c}}{1-\frac{0.825c}{c}}} = 1.70\text{ m}.$$ 

**Discussion**

Because the galaxy is moving away from the Earth, we expect the wavelengths of radiation it emits to be redshifted. The wavelength we calculated is $1.70\text{ m}$, which is redshifted from the original wavelength of $0.525\text{ m}$.

The relativistic Doppler shift is easy to observe. This equation has everyday applications ranging from Doppler-shifted radar velocity measurements of transportation to Doppler-radar storm monitoring. In astronomical observations, the relativistic Doppler shift provides velocity information such as the motion and distance of stars.

**Check Your Understanding**

Suppose a space probe moves away from the Earth at a speed $0.350c$. It sends a radio wave message back to the Earth at a frequency of $1.50\text{ GHz}$. At what frequency is the message received on the Earth?

**Solution**

$$f_{\text{obs}} = f_s \sqrt{\frac{1-\frac{u}{c}}{1+\frac{u}{c}}} = (1.50\text{ GHz}) \sqrt{\frac{1-\frac{0.350c}{c}}{1+\frac{0.350c}{c}}} = 1.04\text{ GHz}$$

## 28.5 Relativistic Momentum

In classical physics, momentum is a simple product of mass and velocity. However, we saw in the last section that when special relativity is taken into account, massive objects have a speed limit. What effect do you think mass and velocity have on the momentum of objects moving at relativistic speeds?

Momentum is one of the most important concepts in physics. The broadest form of Newton’s second law is stated in terms of momentum. Momentum is conserved whenever the net external force on a system is zero. This makes momentum conservation a fundamental tool for analyzing collisions. All of Work, Energy, and Energy Resources is devoted to momentum, and momentum has been important for many other topics as well, particularly where collisions were involved. We will see that momentum has the same importance in modern physics. Relativistic momentum is conserved, and much of what we know about subatomic structure comes from the analysis of collisions of accelerator-produced relativistic particles.

![Figure 28.18](image-url) Momentum is an important concept for these football players from the University of California at Berkeley and the University of California at Davis. Players with more mass often have a larger impact because their momentum is larger. For objects moving at relativistic speeds, the effect is even greater. (credit: John Martinez Pavliga)
The first postulate of relativity states that the laws of physics are the same in all inertial frames. Does the law of conservation of momentum survive this requirement at high velocities? The answer is yes, provided that the momentum is defined as follows.

**Relativistic Momentum**

Relativistic momentum \( p \) is classical momentum multiplied by the relativistic factor \( \gamma \).

\[
p = \gamma mu, \tag{28.40}
\]

where \( m \) is the **rest mass** of the object, \( u \) is its velocity relative to an observer, and the relativistic factor

\[
\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{28.41}
\]

Note that we use \( u \) for velocity here to distinguish it from relative velocity \( v \) between observers. Only one observer is being considered here. With \( p \) defined in this way, total momentum \( p_{\text{tot}} \) is conserved whenever the net external force is zero, just as in classical physics. Again we see that the relativistic quantity becomes virtually the same as the classical at low velocities. That is, relativistic momentum \( \gamma mu \) becomes the classical \( mu \) at low velocities, because \( \gamma \) is very nearly equal to 1 at low velocities.

Relativistic momentum has the same intuitive feel as classical momentum. It is greatest for large masses moving at high velocities, but, because of the factor \( \gamma \), relativistic momentum approaches infinity as \( u \) approaches \( c \). (See Figure 28.19.) This is another indication that an object with mass cannot reach the speed of light. If it did, its momentum would become infinite, an unreasonable value.

![Figure 28.19](http://cnx.org/content/col11406/1.7)

**Misconception Alert: Relativistic Mass and Momentum**

The relativistically correct definition of momentum as \( p = \gamma mu \) is sometimes taken to imply that mass varies with velocity: \( m_{\text{var}} = \gamma m \), particularly in older textbooks. However, note that \( m \) is the mass of the object as measured by a person at rest relative to the object. Thus, \( m \) is defined to be the rest mass, which could be measured at rest, perhaps using gravity. When a mass is moving relative to an observer, the only way that its mass can be determined is through collisions or other means in which momentum is involved. Since the mass of a moving object cannot be determined independently of momentum, the only meaningful mass is rest mass. Thus, when we use the term mass, assume it to be identical to rest mass.

Relativistic momentum is defined in such a way that the conservation of momentum will hold in all inertial frames. Whenever the net external force on a system is zero, relativistic momentum is conserved, just as is the case for classical momentum. This has been verified in numerous experiments.

In **Relativistic Energy**, the relationship of relativistic momentum to energy is explored. That subject will produce our first inkling that objects without mass may also have momentum.

**Check Your Understanding**

What is the momentum of an electron traveling at a speed \( 0.985c \)? The rest mass of the electron is \( 9.11 \times 10^{-31} \) kg.

**Solution**

\[
p = \gamma mu = \frac{mu}{\sqrt{1 - \frac{u^2}{c^2}}} = \frac{(9.11 \times 10^{-31} \text{ kg})(0.985)(3.00 \times 10^8 \text{ m/s})}{\sqrt{1 - \frac{(0.985c)^2}{c^2}}} = 1.56 \times 10^{-21} \text{ kg} \cdot \text{m/s} \tag{28.42}
\]
28.6 Relativistic Energy

A tokamak is a form of experimental fusion reactor, which can change mass to energy. Accomplishing this requires an understanding of relativistic energy. Nuclear reactors are proof of the conservation of relativistic energy.

Conservation of energy is one of the most important laws in physics. Not only does energy have many important forms, but each form can be converted to any other. We know that classically the total amount of energy in a system remains constant. Relativistically, energy is still conserved, provided its definition is altered to include the possibility of mass changing to energy, as in the reactions that occur within a nuclear reactor. Relativistic energy is intentionally defined so that it will be conserved in all inertial frames, just as is the case for relativistic momentum. As a consequence, we learn that several fundamental quantities are related in ways not known in classical physics. All of these relationships are verified by experiment and have fundamental consequences. The altered definition of energy contains some of the most fundamental and spectacular new insights into nature found in recent history.

Total Energy and Rest Energy

The first postulate of relativity states that the laws of physics are the same in all inertial frames. Einstein showed that the law of conservation of energy is valid relativistically, if we define energy to include a relativistic factor.

Total Energy

The total energy is defined as

\[ E = 
\]

\[ = \gamma mc^2, \]  

(28.43)

where \( m \) is mass, \( c \) is the speed of light, \( \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \), and \( v \) is the velocity of the mass relative to an observer. There are many aspects of the total energy \( E \) that we will discuss—among them are how kinetic and potential energies are included in \( E \), and how \( E \) is related to relativistic momentum. But first, note that at rest, total energy is not zero. Rather, when \( v = 0 \), we have \( \gamma = 1 \), and an object has rest energy.

Rest Energy

The rest energy is

\[ E_0 = mc^2. \]  

(28.44)

This is the correct form of Einstein’s most famous equation, which for the first time showed that energy is related to the mass of an object at rest. For example, if energy is stored in the object, its rest mass increases. This also implies that mass can be destroyed to release energy. The implications of these first two equations regarding relativistic energy are so broad that they were not completely recognized for some years after Einstein published them in 1907, nor was the experimental proof that they are correct widely recognized at first. Einstein, it should be noted, did understand and describe the meanings and implications of his theory.

Example 28.6 Calculating Rest Energy: Rest Energy is Very Large

Calculate the rest energy of a 1.00-g mass.

**Strategy**

One gram is a small mass—less than half the mass of a penny. We can multiply this mass, in SI units, by the speed of light squared to find the equivalent rest energy.

**Solution**

1. Identify the knowns. \( m = 1.00 \times 10^{-3} \) kg; \( c = 3.00 \times 10^8 \) m/s

\[ E_0 = mc^2 = (1.00 \times 10^{-3}) \times (3.00 \times 10^8)^2 \]

\[ = 2.7 \times 10^{13} \text{ J} \]
2. Identify the unknown. \( E_0 \)

3. Choose the appropriate equation. \( E_0 = mc^2 \)

4. Plug the knowns into the equation.
   \[
   E_0 = mc^2 = (1.00 \times 10^{-3} \text{ kg})(3.00 \times 10^8 \text{ m/s})^2
   = 9.00 \times 10^{13} \text{ kg} \cdot \text{m}^2/\text{s}^2
   \]

5. Convert units.
   Noting that \( 1 \text{ kg} \cdot \text{m}^2/\text{s}^2 = 1 \text{ J} \), we see the rest mass energy is
   \[
   E_0 = 9.00 \times 10^{13} \text{ J}.
   \]

**Discussion**

This is an enormous amount of energy for a 1.00-g mass. We do not notice this energy, because it is generally not available. Rest energy is large because the speed of light \( c \) is a large number and \( c^2 \) is a very large number, so that \( mc^2 \) is huge for any macroscopic mass. The \( 9.00 \times 10^{13} \text{ J} \) rest mass energy for 1.00 g is about twice the energy released by the Hiroshima atomic bomb and about 10,000 times the kinetic energy of a large aircraft carrier. If a way can be found to convert rest mass energy into some other form (and all forms of energy can be converted into one another), then huge amounts of energy can be obtained from the destruction of mass.

Today, the practical applications of the **conversion of mass into another form of energy**, such as in nuclear weapons and nuclear power plants, are well known. But examples also existed when Einstein first proposed the correct form of relativistic energy, and he did describe some of them. Nuclear radiation had been discovered in the previous decade, and it had been a mystery as to where its energy originated. The explanation was that, in certain nuclear processes, a small amount of mass is destroyed and energy is released and carried by nuclear radiation. But the amount of mass destroyed is so small that it is difficult to detect that any is missing. Although Einstein proposed this as the source of energy in the radioactive salts then being studied, it was many years before there was broad recognition that mass could be and, in fact, commonly is converted to energy. (See Figure 28.21.)

**Figure 28.21** The Sun (a) and the Susquehanna Steam Electric Station (b) both convert mass into energy—the Sun via nuclear fusion, the electric station via nuclear fission. (credits: (a) NASA/Goddard Space Flight Center, Scientific Visualization Studio; (b) U.S. government)

Because of the relationship of rest energy to mass, we now consider mass to be a form of energy rather than something separate. There had not even been a hint of this prior to Einstein's work. Such conversion is now known to be the source of the Sun's energy, the energy of nuclear decay, and even the source of energy keeping Earth's interior hot.

**Stored Energy and Potential Energy**

What happens to energy stored in an object at rest, such as the energy put into a battery by charging it, or the energy stored in a toy gun's compressed spring? The energy input becomes part of the total energy of the object and, thus, increases its rest mass. All stored and potential energy becomes mass in a system. Why is it we don't ordinarily notice this? In fact, conservation of mass (meaning total mass is constant) was one of the great laws verified by 19th-century science. Why was it not noticed to be incorrect? The following example helps answer these questions.
A car battery is rated to be able to move 600 ampere-hours (A·h) of charge at 12.0 V. (a) Calculate the increase in rest mass of such a battery when it is taken from being fully depleted to being fully charged. (b) What percent increase is this, given the battery’s mass is 20.0 kg?

Strategy
In part (a), we first must find the energy stored in the battery, which equals what the battery can supply in the form of electrical potential energy. Since $PE_{elec} = qV$, we have to calculate the charge $q$ in 600 A·h, which is the product of the current $I$ and the time $t$. We then multiply the result by 12.0 V. We can then calculate the battery’s increase in mass using $\Delta E = PE_{elec} = (\Delta m)c^2$. Part (b) is a simple ratio converted to a percentage.

Solution for (a)

1. Identify the knowns. $I \cdot t = 600 \text{ A} \cdot \text{h}$; $V = 12.0 \text{ V}$; $c = 3.00 \times 10^8 \text{ m/s}$
2. Identify the unknown. $\Delta m$
3. Choose the appropriate equation. $PE_{elec} = (\Delta m)c^2$
4. Rearrange the equation to solve for the unknown. $\Delta m = \frac{PE_{elec}}{c^2}$
5. Plug the knowns into the equation.

$$\Delta m = \frac{PE_{elec}}{c^2}$$

$$= \frac{qV}{c^2}$$

$$= \frac{(It)V}{c^2}$$

$$= \frac{(600 \text{ A} \cdot \text{h})(12.0 \text{ V})}{(3.00 \times 10^8 \text{ m/s})^2}$$

Write amperes A as coulombs per second (C/s), and convert hours to seconds.

$$\Delta m = \frac{(600 \text{ C/s} \cdot \text{h})(12.0 \text{ J/C})}{(3.00 \times 10^8 \text{ m/s})^2}$$

$$= \frac{(2.16 \times 10^6 \text{ C})(12.0 \text{ J/C})}{(3.00 \times 10^8 \text{ m/s})^2}$$

Using the conversion $1 \text{ kg} \cdot \text{m}^2/\text{s}^2 = 1 \text{ J}$, we can write the mass as

$$\Delta m = 2.88 \times 10^{-10} \text{ kg}.$$

Solution for (b)

1. Identify the knowns. $\Delta m = 2.88 \times 10^{-10} \text{ kg}$; $m = 20.0 \text{ kg}$
2. Identify the unknown. % change
3. Choose the appropriate equation. $\% \text{ increase} = \frac{\Delta m}{m} \times 100\%$
4. Plug the knowns into the equation.

$$\% \text{ increase} = \frac{\Delta m}{m} \times 100\%$$

$$= \frac{2.88 \times 10^{-10} \text{ kg}}{20.0 \text{ kg}} \times 100\%$$

$$= 1.44 \times 10^{-9} \%.$$

Discussion
Both the actual increase in mass and the percent increase are very small, since energy is divided by $c^2$, a very large number. We would have to be able to measure the mass of the battery to a precision of a billionth of a percent, or 1 part in $10^{11}$, to notice this increase. It is no wonder that the mass variation is not readily observed. In fact, this change in mass is so small that we may question how you could verify it is real. The answer is found in nuclear processes in which the percentage of mass destroyed is large enough to be measured. The mass of the fuel of a nuclear reactor, for example, is measurably smaller when its energy has been used. In that case, stored energy has been released (converted mostly to heat and electricity) and the rest mass has decreased. This is also the case when you use the energy stored in a battery, except that the stored energy is much greater in nuclear processes, making the change in mass measurable in practice as well as in theory.
Kinetic Energy and the Ultimate Speed Limit

Kinetic energy is energy of motion. Classically, kinetic energy has the familiar expression \( \frac{1}{2}mv^2 \). The relativistic expression for kinetic energy is obtained from the work-energy theorem. This theorem states that the net work on a system goes into kinetic energy. If our system starts from rest, then the work-energy theorem is

\[
W_{\text{net}} = KE. \tag{28.50}
\]

Relativistically, at rest we have rest energy \( E_0 = mc^2 \). The work increases this to the total energy \( E = \gamma mc^2 \). Thus,

\[
W_{\text{net}} = E - E_0 = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2. \tag{28.51}
\]

Relativistically, we have \( W_{\text{net}} = KE_{\text{rel}} \).

**Relativistic Kinetic Energy**

Relativistic kinetic energy is

\[
KE_{\text{rel}} = (\gamma - 1)mc^2. \tag{28.52}
\]

When motionless, we have \( v = 0 \) and

\[
\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1, \tag{28.53}
\]

so that \( KE_{\text{rel}} = 0 \) at rest, as expected. But the expression for relativistic kinetic energy (such as total energy and rest energy) does not look much like the classical \( \frac{1}{2}mv^2 \). To show that the classical expression for kinetic energy is obtained at low velocities, we note that the binomial expansion for \( \gamma \) at low velocities gives

\[
\gamma = 1 + \frac{1}{2} \frac{v^2}{c^2}.
\]

A binomial expansion is a way of expressing an algebraic quantity as a sum of an infinite series of terms. In some cases, as in the limit of small velocity here, most terms are very small. Thus the expression derived for \( \gamma \) here is not exact, but it is a very accurate approximation. Thus, at low velocities,

\[
\gamma - 1 = \frac{1}{2} \frac{v^2}{c^2}. \tag{28.55}
\]

Entering this into the expression for relativistic kinetic energy gives

\[
KE_{\text{rel}} = \left[ 1 - \frac{1}{2} \frac{v^2}{c^2} \right] mc^2 = \frac{1}{2}mv^2 = KE_{\text{class}}. \tag{28.56}
\]

So, in fact, relativistic kinetic energy does become the same as classical kinetic energy when \( v << c \).

It is even more interesting to investigate what happens to kinetic energy when the velocity of an object approaches the speed of light. We know that \( \gamma \) becomes infinite as \( v \) approaches \( c \), so that \( KE_{\text{rel}} \) also becomes infinite as the velocity approaches the speed of light. (See Figure 28.22.) An infinite amount of work (and, hence, an infinite amount of energy input) is required to accelerate a mass to the speed of light.

**The Speed of Light**

No object with mass can attain the speed of light.

So the speed of light is the ultimate speed limit for any particle having mass. All of this is consistent with the fact that velocities less than \( c \) always add to less than \( c \). Both the relativistic form for kinetic energy and the ultimate speed limit being \( c \) have been confirmed in detail in numerous experiments. No matter how much energy is put into accelerating a mass, its velocity can only approach—not reach—the speed of light.
This graph of $KE_{\text{rel}}$ versus velocity shows how kinetic energy approaches infinity as velocity approaches the speed of light. It is thus not possible for an object having mass to reach the speed of light. Also shown is $KE_{\text{class}}$, the classical kinetic energy, which is similar to relativistic kinetic energy at low velocities. Note that much more energy is required to reach high velocities than predicted classically.

**Example 28.8 Comparing Kinetic Energy: Relativistic Energy Versus Classical Kinetic Energy**

An electron has a velocity $v = 0.990c$. (a) Calculate the kinetic energy in MeV of the electron. (b) Compare this with the classical value for kinetic energy at this velocity. (The mass of an electron is $9.11 \times 10^{-31}$ kg.)

**Strategy**
The expression for relativistic kinetic energy is always correct, but for (a) it must be used since the velocity is highly relativistic (close to $c$). First, we will calculate the relativistic factor $\gamma$, and then use it to determine the relativistic kinetic energy. For (b), we will calculate the classical kinetic energy (which would be close to the relativistic value if $v$ were less than a few percent of $c$) and see that it is not the same.

**Solution for (a)**

1. Identify the knowns. $v = 0.990c$; $m = 9.11 \times 10^{-31}$ kg
2. Identify the unknown. $KE_{\text{rel}}$
3. Choose the appropriate equation. $KE_{\text{rel}} = (\gamma - 1)mc^2$
4. Plug the knowns into the equation.
   First calculate $\gamma$. We will carry extra digits because this is an intermediate calculation.
   \[
   \gamma = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}} = \frac{1}{1 - (0.990)^2} = 7.0888
   \]
   Next, we use this value to calculate the kinetic energy.
   \[
   KE_{\text{rel}} = (\gamma - 1)mc^2
   = (7.0888 - 1)(9.11 \times 10^{-31} \text{ kg})(3.00 \times 10^8 \text{ m/s})^2
   = 4.99 \times 10^{-13} \text{ J}
   \]
5. Convert units.
   \[
   KE_{\text{rel}} = (4.99 \times 10^{-13} \text{ J})\left(\frac{1 \text{ MeV}}{1.60 \times 10^{-13} \text{ J}}\right)
   = 3.12 \text{ MeV}
   \]

**Solution for (b)**

1. List the knowns. $v = 0.990c$; $m = 9.11 \times 10^{-31}$ kg
2. List the unknown. $KE_{\text{class}}$
3. Choose the appropriate equation. $KE_{\text{class}} = \frac{1}{2}mv^2$
4. Plug the knowns into the equation.
\[ KE_{\text{class}} = \left( \frac{1}{2} \right) mv^2 \]
\[ = \left( \frac{1}{2} \right)(9.00 \times 10^{-31} \text{ kg})(0.990)(3.00 \times 10^8 \text{ m/s})^2 \]
\[ = 4.02 \times 10^{-14} \text{ J} \]

5. Convert units.

\[ KE_{\text{class}} = 4.02 \times 10^{-14} \text{ J} \left( \frac{1 \text{ MeV}}{1.60 \times 10^{-13} \text{ J}} \right) \]
\[ = 0.251 \text{ MeV} \]

Discussion

As might be expected, since the velocity is 99.0% of the speed of light, the classical kinetic energy is significantly off from the correct relativistic value. Note also that the classical value is much smaller than the relativistic value. In fact, \( KE_{\text{rel}}/KE_{\text{class}} = 12.4 \) here. This is some indication of how difficult it is to get a mass moving close to the speed of light. Much more energy is required than predicted classically. Some people interpret this extra energy as going into increasing the mass of the system, but, as discussed in Relativistic Momentum, this cannot be verified unambiguously. What is certain is that ever-increasing amounts of energy are needed to get the velocity of a mass a little closer to that of light. An energy of 3 MeV is a very small amount for an electron, and it can be achieved with present-day particle accelerators. SLAC, for example, can accelerate electrons to over \( 50 \times 10^9 \text{ eV} = 50,000 \text{ MeV} \).

Is there any point in getting \( v \) a little closer to \( c \) than 99.0% or 99.9%? The answer is yes. We learn a great deal by doing this. The energy that goes into a high-velocity mass can be converted to any other form, including into entirely new masses. (See Figure 28.23.) Most of what we know about the substructure of matter and the collection of exotic short-lived particles in nature has been learned this way. Particles are accelerated to extremely relativistic energies and made to collide with other particles, producing totally new species of particles. Patterns in the characteristics of these previously unknown particles hint at a basic substructure for all matter. These particles and some of their characteristics will be covered in Particle Physics.

![Figure 28.23](credit: Fermilab, Reidar Hahn)

Relativistic Energy and Momentum

We know classically that kinetic energy and momentum are related to each other, since

\[ KE_{\text{class}} = \left( \frac{1}{2} \right) mv^2 = \left( \frac{1}{2} \right) (mv)^2 = \frac{1}{2} mv^2. \]

Relativistically, we can obtain a relationship between energy and momentum by algebraically manipulating their definitions. This produces

\[ E^2 = (pc)^2 + (mc^2)^2, \]

where \( E \) is the relativistic total energy and \( p \) is the relativistic momentum. This relationship between relativistic energy and relativistic momentum is more complicated than the classical, but we can gain some interesting new insights by examining it. First, total energy is related to momentum and rest mass. At rest, momentum is zero, and the equation gives the total energy to be the rest energy \( mc^2 \) (so this equation is consistent with the discussion of rest energy above). However, as the mass is accelerated, its momentum \( p \) increases, thus increasing the total energy. At sufficiently high velocities, the rest energy term \( (mc^2)^2 \) becomes negligible compared with the momentum term \( (pc)^2 \); thus, \( E = pc \) at extremely relativistic velocities.

If we consider momentum \( p \) to be distinct from mass, we can determine the implications of the equation \( E^2 = (pc)^2 + (mc^2)^2 \), for a particle that has no mass. If we take \( m \) to be zero in this equation, then \( E = pc \), or \( p = E/c \). Massless particles have this momentum. There are several massless particles found in nature, including photons (these are quanta of electromagnetic radiation). Another implication is that a massless particle must travel at speed \( c \) and only at speed \( c \). While it is beyond the scope of this text to examine the relationship in the equation \( E^2 = (pc)^2 + (mc^2)^2 \), in detail, we can see that the relationship has important implications in special relativity.
Problem-Solving Strategies for Relativity

1. Examine the situation to determine that it is necessary to use relativity. Relativistic effects are related to \( \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \), the quantitative relativistic factor. If \( \gamma \) is very close to 1, then relativistic effects are small and differ very little from the usually easier classical calculations.

2. Identify exactly what needs to be determined in the problem (identify the unknowns).

3. Make a list of what is given or can be inferred from the problem as stated (identify the knowns). Look in particular for information on relative velocity \( v \).

4. Make certain you understand the conceptual aspects of the problem before making any calculations. Decide, for example, which observer sees time dilated or length contracted before plugging into equations. If you have thought about who sees what, who is moving with the event being observed, who sees proper time, and so on, you will find it much easier to determine if your calculation is reasonable.

5. Determine the primary type of calculation to be done to find the unknowns identified above. You will find the section summary helpful in determining whether a length contraction, relativistic kinetic energy, or some other concept is involved.

6. Do not round off during the calculation. As noted in the text, you must often perform your calculations to many digits to see the desired effect. You may round off at the very end of the problem, but do not use a rounded number in a subsequent calculation.

7. Check the answer to see if it is reasonable: Does it make sense? This may be more difficult for relativity, since we do not encounter it directly. But you can look for velocities greater than \( c \) or relativistic effects that are in the wrong direction (such as a time contraction where a dilation was expected).

Check Your Understanding

A photon decays into an electron-positron pair. What is the kinetic energy of the electron if its speed is \( 0.992c \)?

Solution

\[
\begin{align*}
\text{KE}_{\text{rel}} &= (\gamma - 1)mc^2 = \left( \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right)mc^2 \\
&= \left( \frac{1}{\sqrt{1 - \left(0.992c\right)^2}} - 1 \right)(9.11 \times 10^{-31} \text{ kg})(3.00 \times 10^8 \text{ m/s})^2 = 5.67 \times 10^{-13} \text{ J}
\end{align*}
\] (28.64)

Glossary

classical velocity addition: the method of adding velocities when \( v \ll c \); velocities add like regular numbers in one-dimensional motion:

\[
u = v + u'
\]

where \( v \) is the velocity between two observers, \( u \) is the velocity of an object relative to one observer, and \( u' \) is the velocity relative to the other observer

first postulate of special relativity: the idea that the laws of physics are the same and can be stated in their simplest form in all inertial frames of reference

inertial frame of reference: a reference frame in which a body at rest remains at rest and a body in motion moves at a constant speed in a straight line unless acted on by an outside force

length contraction:

\[
L = L_0 \sqrt{1 - \frac{v^2}{c^2}} = \frac{L_0}{\gamma}
\]

Michelson-Morley experiment: an investigation performed in 1887 that proved that the speed of light in a vacuum is the same in all frames of reference from which it is viewed

proper length: \( L_0 \); the distance between two points measured by an observer who is at rest relative to both of the points; Earth-bound observers measure proper length when measuring the distance between two points that are stationary relative to the Earth

proper time:

\[
\Delta t_0 \text{, the time measured by an observer at rest relative to the event being observed: } \Delta t = \frac{\Delta t_0}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma \Delta t_0 \text{, where } \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}
\]

relativistic Doppler effects: a change in wavelength of radiation that is moving relative to the observer; the wavelength of the radiation is longer (called a red shift) than that emitted by the source when the source moves away from the observer and shorter (called a blue shift) when the source moves toward the observer; the shifted wavelength is described by the equation

\[
\lambda_{\text{obs}} = \lambda_s \sqrt{1 + \frac{u}{c}}
\]

where \( \lambda_{\text{obs}} \) is the observed wavelength, \( \lambda_s \) is the source wavelength, and \( u \) is the velocity of the source to the observer
**relativistic kinetic energy:** the kinetic energy of an object moving at relativistic speeds: \( KE_{\text{rel}} = (\gamma - 1)mc^2 \), where \( \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \)

**relativistic momentum:** \( p \), the momentum of an object moving at relativistic velocity; \( p = \gamma mu \), where \( m \) is the rest mass of the object, \( u \) is its velocity relative to an observer, and the relativistic factor \( \gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \)

**relativistic velocity addition:** the method of adding velocities of an object moving at a relativistic speed: \( u = \frac{v + u'}{1 + \frac{vu'}{c^2}} \), where \( v \) is the relative velocity between two observers, \( u \) is the velocity of an object relative to one observer, and \( u' \) is the velocity relative to the other observer

**relativity:** the study of how different observers measure the same event

**rest energy:** the energy stored in an object at rest: \( E_0 = mc^2 \)

**rest mass:** the mass of an object as measured by a person at rest relative to the object

**second postulate of special relativity:** the idea that the speed of light \( c \) is a constant, independent of the relative motion of the source and observer

**special relativity:** the theory that, in an inertial frame of reference, the motion of an object is relative to the frame from which it is viewed or measured

**time dilation:** the phenomenon of time passing slower to an observer who is moving relative to another observer

**total energy:** defined as \( E = \gamma mc^2 \), where \( \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \)

**twin paradox:** this asks why a twin traveling at a relativistic speed away and then back towards the Earth ages less than the Earth-bound twin. The premise to the paradox is faulty because the traveling twin is accelerating, and special relativity does not apply to accelerating frames of reference

---

**Section Summary**

28.1 Einstein’s Postulates

- Relativity is the study of how different observers measure the same event.
- Modern relativity is divided into two parts. Special relativity deals with observers who are in uniform (unaccelerated) motion, whereas general relativity includes accelerated relative motion and gravity. Modern relativity is correct in all circumstances and, in the limit of low velocity and weak gravitation, gives the same predictions as classical relativity.
- An inertial frame of reference is a reference frame in which a body at rest remains at rest and a body in motion moves at a constant speed in a straight line unless acted on by an outside force.
- Modern relativity is based on Einstein’s two postulates. The first postulate of special relativity is the idea that the laws of physics are the same and can be stated in their simplest form in all inertial frames of reference. The second postulate of special relativity is the idea that the speed of light \( c \) is a constant, independent of the relative motion of the source and observer.
- The Michelson-Morley experiment demonstrated that the speed of light in a vacuum is independent of the motion of the Earth about the Sun.

28.2 Simultaneity And Time Dilation

- Two events are defined to be simultaneous if an observer measures them as occurring at the same time. They are not necessarily simultaneous to all observers—simultaneity is not absolute.
- Time dilation is the phenomenon of time passing slower for an observer who is moving relative to another observer.
- Observers moving at a relative velocity \( \nu \) do not measure the same elapsed time for an event. Proper time \( \Delta t_0 \) is the time measured by an observer at rest relative to the event being observed. Proper time is related to the time \( \Delta t \) measured by an Earth-bound observer by the equation

\[
\Delta t = \frac{\Delta t_0}{\sqrt{1 - \frac{\nu^2}{c^2}}} = \gamma \Delta t_0
\]

where

\[
\gamma = \frac{1}{\sqrt{1 - \frac{\nu^2}{c^2}}}
\]

- The equation relating proper time and time measured by an Earth-bound observer implies that relative velocity cannot exceed the speed of light.
• The twin paradox asks why a twin traveling at a relativistic speed away and then back towards the Earth ages less than the Earth-bound twin. The premise to the paradox is faulty because the traveling twin is accelerating. Special relativity does not apply to accelerating frames of reference.
• Time dilation is usually negligible at low relative velocities, but it does occur, and it has been verified by experiment.

28.3 Length Contraction
• All observers agree upon relative speed.
• Distance depends on an observer’s motion. Proper length $L_0$ is the distance between two points measured by an observer who is at rest relative to both of the points. Earth-bound observers measure proper length when measuring the distance between two points that are stationary relative to the Earth.
• Length contraction $L$ is the shortening of the measured length of an object moving relative to the observer’s frame:

$$L = L_0 \sqrt{1 - \frac{v^2}{c^2}} = \frac{L_0}{\gamma}.$$ 

28.4 Relativistic Addition of Velocities
• With classical velocity addition, velocities add like regular numbers in one-dimensional motion: $u = v + u'$, where $v$ is the velocity between two observers, $u$ is the velocity of an object relative to one observer, and $u'$ is the velocity relative to the other observer.
• Velocities cannot add to be greater than the speed of light. Relativistic velocity addition describes the velocities of an object moving at a relativistic speed:

$$u = \frac{v + u'}{1 + \frac{vu'}{c^2}}.$$ 

An observer of electromagnetic radiation sees relativistic Doppler effects if the source of the radiation is moving relative to the observer. The wavelength of the radiation is longer (called a red shift) than that emitted by the source when the source moves away from the observer and shorter (called a blue shift) when the source moves toward the observer. The shifted wavelength is described by the equation

$$\lambda_{obs} = \lambda_s \sqrt{1 - \frac{u}{c}}.$$ 

$\lambda_{obs}$ is the observed wavelength, $\lambda_s$ is the source wavelength, and $u$ is the relative velocity of the source to the observer.

28.5 Relativistic Momentum
• The law of conservation of momentum is valid whenever the net external force is zero and for relativistic momentum. Relativistic momentum $p$ is classical momentum multiplied by the relativistic factor $\gamma$.

$$p = \gamma mu,$$

where $m$ is the rest mass of the object, $u$ is its velocity relative to an observer, and the relativistic factor $\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$.

At low velocities, relativistic momentum is equivalent to classical momentum.
• Relativistic momentum approaches infinity as $u$ approaches $c$. This implies that an object with mass cannot reach the speed of light.
• Relativistic momentum is conserved, just as classical momentum is conserved.

28.6 Relativistic Energy
• Relativistic energy is conserved as long as we define it to include the possibility of mass changing to energy.
• Total Energy is defined as: $E = \gamma mc^2$, where $\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$. Rest energy is $E_0 = mc^2$, meaning that mass is a form of energy. If energy is stored in an object, its mass increases. Mass can be destroyed to release energy.
• We do not ordinarily notice the increase or decrease in mass of an object because the change in mass is so small for a large increase in energy.
• The relativistic work-energy theorem is $W_{net} = E - E_0 = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2$.
• Relativistically, $W_{net} = KE_{rel}$, where $KE_{rel}$ is the relativistic kinetic energy.

Relativistic kinetic energy is $KE_{rel} = (\gamma - 1)mc^2$, where $\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$. At low velocities, relativistic kinetic energy reduces to classical kinetic energy.

• No object with mass can attain the speed of light because an infinite amount of work and an infinite amount of energy input is required to accelerate a mass to the speed of light.
• The equation $E^2 = (pc)^2 + (mc)^2$ relates the relativistic total energy $E$ and the relativistic momentum $p$. At extremely high velocities, the rest energy $mc^2$ becomes negligible, and $E = pc$. 

CHAPTER 28 | SPECIAL RELATIVITY 1023
# Conceptual Questions

## 28.1 Einstein's Postulates

1. Which of Einstein's postulates of special relativity includes a concept that does not fit with the ideas of classical physics? Explain.


3. When you are flying in a commercial jet, it may appear to you that the airplane is stationary and the Earth is moving beneath you. Is this point of view valid? Discuss briefly.

## 28.2 Simultaneity And Time Dilation

4. Does motion affect the rate of a clock as measured by an observer moving with it? Does motion affect how an observer moving relative to a clock measures its rate?

5. To whom does the elapsed time for a process seem to be longer, an observer moving relative to the process or an observer moving with the process? Which observer measures proper time?

6. How could you travel far into the future without aging significantly? Could this method also allow you to travel into the past?

## 28.3 Length Contraction

7. To whom does an object seem greater in length, an observer moving with the object or an observer moving relative to the object? Which observer measures the object's proper length?

8. Relativistic effects such as time dilation and length contraction are present for cars and airplanes. Why do these effects seem strange to us?

9. Suppose an astronaut is moving relative to the Earth at a significant fraction of the speed of light. (a) Does he observe the rate of his clocks to have slowed? (b) What change in the rate of Earth-bound clocks does he see? (c) Does his ship seem to him to shorten? (d) What about the distance between stars that lie on lines parallel to his motion? (e) Do he and an Earth-bound observer agree on his velocity relative to the Earth?

## 28.4 Relativistic Addition of Velocities

10. Explain the meaning of the terms “red shift” and “blue shift” as they relate to the relativistic Doppler effect.

11. What happens to the relativistic Doppler effect when relative velocity is zero? Is this the expected result?

12. Is the relativistic Doppler effect consistent with the classical Doppler effect in the respect that $\lambda_{\text{obs}}$ is larger for motion away?

13. All galaxies farther away than about $50 \times 10^6$ ly exhibit a red shift in their emitted light that is proportional to distance, with those farther and farther away having progressively greater red shifts. What does this imply, assuming that the only source of red shift is relative motion? (Hint: At these large distances, it is space itself that is expanding, but the effect on light is the same.)

## 28.5 Relativistic Momentum

14. How does modern relativity modify the law of conservation of momentum?

15. Is it possible for an external force to be acting on a system and relativistic momentum to be conserved? Explain.

## 28.6 Relativistic Energy

16. How are the classical laws of conservation of energy and conservation of mass modified by modern relativity?

17. What happens to the mass of water in a pot when it cools, assuming no molecules escape or are added? Is this observable in practice? Explain.

18. Consider a thought experiment. You place an expanded balloon of air on weighing scales outside in the early morning. The balloon stays on the scales and you are able to measure changes in its mass. Does the mass of the balloon change as the day progresses? Discuss the difficulties in carrying out this experiment.

19. The mass of the fuel in a nuclear reactor decreases by an observable amount as it puts out energy. Is the same true for the coal and oxygen combined in a conventional power plant? If so, is this observable in practice for the coal and oxygen? Explain.

20. We know that the velocity of an object with mass has an upper limit of $c$. Is there an upper limit on its momentum? Its energy? Explain.

21. Given the fact that light travels at $c$, can it have mass? Explain.

22. If you use an Earth-based telescope to project a laser beam onto the Moon, you can move the spot across the Moon’s surface at a velocity greater than the speed of light. Does this violate modern relativity? (Note that light is being sent from the Earth to the Moon, not across the surface of the Moon.)
28.2 Simultaneity And Time Dilation

1. (a) What is \( \gamma \) if \( v = 0.250c \)? (b) If \( v = 0.500c \)?
2. (a) What is \( \gamma \) if \( v = 0.100c \)? (b) If \( v = 0.900c \)?
3. Particles called \( \pi \)-mesons are produced by accelerator beams. If these particles travel at 2.70x10^8 m/s and live 2.60x10^{-8} s when at rest relative to an observer, how long do they live as viewed in the laboratory?

4. Suppose a particle called a kaon is created by cosmic radiation striking the atmosphere. It moves by you at 0.980c, and it lives 1.24x10^{-8} s when at rest relative to an observer. How long does it live as you observe it?

5. A neutral \( \pi \)-meson is a particle that can be created by accelerator beams. If one such particle lives 1.40x10^{-16} s as measured in the laboratory, and 0.840x10^{-16} s when at rest relative to an observer, what is its velocity relative to the laboratory?

6. A neutron lives 900 s when at rest relative to an observer. How fast is the neutron moving relative to an observer who measures its life span to be 2065 s?

7. If relativistic effects are to be less than 1%, then \( \gamma \) must be less than 1.01. At what relative velocity is \( \gamma = 1.01 \)?
8. If relativistic effects are to be less than 3%, then \( \gamma \) must be less than 1.03. At what relative velocity is \( \gamma = 1.03 \)?
9. (a) At what relative velocity is \( \gamma = 1.50 \)? (b) At what relative velocity is \( \gamma = 100 \)?
10. (a) At what relative velocity is \( \gamma = 2.00 \)? (b) At what relative velocity is \( \gamma = 10.0 \)?

11. Unreasonable Results
   (a) Find the value of \( \gamma \) for the following situation. An Earth-bound observer measures 23.9 h to have passed while signals from a high-speed space probe indicate that 24.0 h have passed on board. (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

28.3 Length Contraction

12. A spaceship, 200 m long as seen on board, moves by the Earth at 0.970c. What is its length as measured by an Earth-bound observer?
13. How fast would a 6.0 m-long sports car have to be going past you in order for it to appear only 5.5 m long?
14. (a) How far does the muon in Example 28.1 travel according to the Earth-bound observer? (b) How far does it travel as viewed by an observer moving with it? Base your calculation on its velocity relative to the Earth and the time it lives (proper time). (c) Verify that these two distances are related through length contraction \( \gamma=3.20 \).
15. (a) How long would the muon in Example 28.1 have lived as observed on the Earth if its velocity was 0.0500c? (b) How far would it have traveled as observed on the Earth? (c) What distance is this in the muon’s frame?
16. (a) How long does it take the astronaut in Example 28.2 to travel 4.30 ly at 0.99944c (as measured by the Earth-bound observer)? (b) How long does it take according to the astronaut? (c) Verify that these two times are related through time dilation with \( \gamma=30.00 \) as given.

17. (a) How fast would an athlete need to be running for a 100-m race to look 100 yd long? (b) Is the answer consistent with the fact that relativistic effects are difficult to observe in ordinary circumstances? Explain.

18. Unreasonable Results
   (a) Find the value of \( \gamma \) for the following situation. An astronaut measures the length of her spaceship to be 25.0 m, while an Earth-bound observer measures it to be 100 m. (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

19. Unreasonable Results
   A spaceship is heading directly toward the Earth at a velocity of 0.800c. The astronaut on board claims that he can send a canister toward the Earth at 1.20c relative to the Earth. (a) Calculate the velocity the canister must have relative to the spaceship. (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

28.4 Relativistic Addition of Velocities

20. Suppose a spaceship heading straight towards the Earth at 0.750c can shoot a canister at 0.500c relative to the ship. (a) What is the velocity of the canister relative to the Earth, if it is shot directly at the Earth? (b) If it is shot directly away from the Earth?
21. Repeat the previous problem with the ship heading directly away from the Earth.
22. If a spaceship is approaching the Earth at 0.100c and a message capsule is sent toward it at 0.100c relative to the Earth, what is the speed of the capsule relative to the ship?
23. (a) Suppose the speed of light were only 3000 m/s. A jet fighter moving toward a target on the ground at 800 m/s shoots bullets, each having a muzzle velocity of 1000 m/s. What are the bullets’ velocity relative to the target? (b) If the speed of light was this small, would you observe relativistic effects in everyday life? Discuss.
24. If a galaxy moving away from the Earth has a speed of 1000 km/s and emits 656 nm light characteristic of hydrogen (the most common element in the universe). (a) What wavelength would we observe on the Earth? (b) What type of electromagnetic radiation is this? (c) Why is the speed of the Earth in its orbit negligible here?
25. A space probe speeding towards the nearest star moves at 0.250c and sends radio information at a broadcast frequency of 1.00 GHz. What frequency is received on the Earth?
26. If two spaceships are heading directly towards each other at 0.800c, at what speed must a canister be shot from the first ship to approach the other at 0.999c as seen by the second ship?
27. Two planets are on a collision course, heading directly towards each other at 0.250c. A spaceship sent from one planet approaches the second at 0.750c as seen by the second planet. What is the velocity of the ship relative to the first planet?
28. When a missile is shot from one spaceship towards another, it leaves the first at 0.950c and approaches the other at 0.750c. What is the relative velocity of the two ships?
29. What is the relative velocity of two spaceships if one fires a missile at the other at 0.750c and the other observes it to approach at 0.950c?
30. Near the center of our galaxy, hydrogen gas is moving directly away from us in its orbit about a black hole. We receive 1900 nm electromagnetic radiation and know that it was 1875 nm when emitted by the hydrogen gas. What is the speed of the gas?
31. A highway patrol officer uses a device that measures the speed of vehicles by bouncing radar off them and measuring the Doppler shift. The outgoing radar has a frequency of 100 GHz and the returning echo has a
frequency 15.0 kHz higher. What is the velocity of the vehicle? Note that there are two Doppler shifts in echoes. Be certain not to round off until the end of the problem, because the effect is small.

32. Prove that for any relative velocity \( v \) between two observers, a beam of light sent from one to the other will approach at speed \( c \) (provided that \( v \) is less than \( c \), of course).

33. Show that for any relative velocity \( v \) between two observers, a beam of light projected by one directly away from the other will move away at the speed of light (provided that \( v \) is less than \( c \), of course).

34. (a) All but the closest galaxies are receding from our own Milky Way Galaxy. If a galaxy 12.0x10^9 ly away is receding from us at 0.900c , at what velocity relative to us must we send an exploratory probe to approach the other galaxy at 0.990c , as measured from that galaxy? (b) How long will it take the probe to reach the other galaxy as measured from the Earth? You may assume that the velocity of the other galaxy remains constant. (c) How long will it then take for a radio signal to be beamed back? (All of this is possible in principle, but not practical.)

28.5 Relativistic Momentum

35. Find the momentum of a helium nucleus having a mass of 6.68x10^{-27} kg that is moving at 0.200c .

36. What is the momentum of an electron traveling at 0.980c ?

37. (a) Find the momentum of a 1.00x10^9 kg asteroid heading towards the Earth at 30.0 km/s . (b) Find the ratio of this momentum to the classical momentum. (Hint: Use the approximation that \( \gamma = 1 + (1/2)v^2/c^2 \) at low velocities.)

38. (a) What is the momentum of a 2000 kg satellite orbiting at 4.00 km/s? (b) Find the ratio of this momentum to the classical momentum. (Hint: Use the approximation that \( \gamma = 1 + (1/2)v^2/c^2 \) at low velocities.)

39. What is the velocity of an electron that has a momentum of 3.04x10^{-21} kg⋅m/s? Note that you must calculate the velocity to at least four digits to see the difference from \( c \).

40. Find the velocity of a proton that has a momentum of 4.48x10^{-19} kg⋅m/s.

41. (a) Calculate the speed of a 1.00-μg particle of dust that has the same momentum as a proton moving at 0.999c . (b) What does the small speed tell us about the mass of a proton compared to even a tiny amount of macroscopic matter?

42. (a) Calculate \( \gamma \) for a proton that has a momentum of 1.00 kg⋅m/s. (b) What is its speed? Such protons form a rare component of cosmic radiation with uncertain origins.

28.6 Relativistic Energy

43. What is the rest energy of an electron, given its mass is 9.11x10^{-31} kg? Give your answer in joules and MeV.

44. Find the rest energy in joules and MeV of a proton, given its mass is 1.67x10^{-27} kg.

45. If the rest energies of a proton and a neutron (the two constituents of nuclei) are 938.3 and 939.6 MeV respectively, what is the difference in their masses in kilograms?

46. The Big Bang that began the universe is estimated to have released 10^{68} J of energy. How many stars could half this energy create, assuming the average star’s mass is 4.00x10^{30} kg?

47. A supernova explosion of a 2.00x10^{31} kg star produces 1.00x10^{44} kg of energy. (a) How many kilograms of mass are converted to energy in the explosion? (b) What is the ratio \( \Delta m/m \) of mass destroyed to the original mass of the star?

48. (a) Using data from Table 7.1, calculate the mass converted to energy by the fission of 1.00 kg of uranium. (b) What is the ratio of mass destroyed to the original mass, \( \Delta m/m \)?

49. (a) Using data from Table 7.1, calculate the amount of mass converted to energy by the fusion of 1.00 kg of hydrogen. (b) What is the ratio of mass destroyed to the original mass, \( \Delta m/m \)? (c) How does this compare with \( \Delta m/m \) for the fission of 1.00 kg of uranium?

50. There is approximately 10^{34} J of energy available from fusion of hydrogen in the world's oceans. (a) If 10^{33} J of this energy were utilized, what would be the decrease in mass of the oceans? (b) How great a volume of water does this correspond to? (c) Comment on whether this is a significant fraction of the total mass of the oceans.

51. A muon has a rest mass energy of 105.7 MeV, and it decays into an electron and a massless particle. (a) If all the lost mass is converted into the electron's kinetic energy, find \( \gamma \) for the electron. (b) What is the electron's velocity?

52. A \( \pi \)-meson is a particle that decays into a muon and a massless particle. The \( \pi \)-meson has a rest mass energy of 139.6 MeV, and the muon has a rest mass energy of 105.7 MeV. Suppose the \( \pi \)-meson is at rest and all of the missing mass goes into the muon's kinetic energy. How fast will the muon move?

53. (a) Calculate the relativistic kinetic energy of a 1000-kg car moving at 30.0 m/s if the speed of light were only 45.0 m/s. (b) Find the ratio of the relativistic kinetic energy to classical.

54. Alpha decay is nuclear decay in which a helium nucleus is emitted. If the helium nucleus has a mass of 6.80x10^{-27} kg and is given 5.00 MeV of kinetic energy, what is its velocity?

55. (a) Beta decay is nuclear decay in which an electron is emitted. If the electron is given 0.750 MeV of kinetic energy, what is its velocity? (b) Comment on how the high velocity is consistent with the kinetic energy as it compares to the rest mass energy of the electron.

56. A positron is an antimatter version of the electron, having exactly the same mass. When a positron and an electron meet, they annihilate, converting all of their mass into energy. (a) Find the energy released, assuming negligible kinetic energy before the annihilation. (b) If this energy is given to a proton in the form of kinetic energy, what is its velocity? (c) If this energy is given to another electron in the form of kinetic energy, what is its velocity?

57. What is the kinetic energy in MeV of a \( \pi \)-meson that lives 1.40x10^{-16} s as measured in the laboratory, and 0.840x10^{-16} s when at rest relative to an observer, given that its rest energy is 135 MeV?

58. Find the kinetic energy in MeV of a neutron with a measured life span of 2065 s, given its rest energy is 939.6 MeV, and rest life span is 900s.

59. (a) Show that \( (pc)^2/(mc)^2 = \gamma^2 - 1 \). This means that at large velocities \( pc \gg mc^2 \). (b) Is \( E \approx pc \) when \( \gamma = 30.0 \), as for the astronaut discussed in the twin paradox?

60. One cosmic ray neutron has a velocity of 0.250c relative to the Earth. (a) What is the neutron's total energy in MeV? (b) Find its momentum. (c) Is \( E \approx pc \) in this situation? Discuss in terms of the equation given in part (a) of the previous problem.
61. What is $\gamma$ for a proton having a mass energy of 938.3 MeV accelerated through an effective potential of 1.0 TV (teravolt) at Fermilab outside Chicago?

62. (a) What is the effective accelerating potential for electrons at the Stanford Linear Accelerator, if $\gamma = 1.00 \times 10^5$ for them? (b) What is their total energy (nearly the same as kinetic in this case) in GeV?

63. (a) Using data from Table 7.1, find the mass destroyed when the energy in a barrel of crude oil is released. (b) Given these barrels contain 200 liters and assuming the density of crude oil is $750 \text{ kg/m}^3$, what is the ratio of mass destroyed to original mass, $\Delta m / m$?

64. (a) Calculate the energy released by the destruction of 1.00 kg of mass. (b) How many kilograms could be lifted to a 10.0 km height by this amount of energy?

65. A Van de Graaff accelerator utilizes a 50.0 MV potential difference to accelerate charged particles such as protons. (a) What is the velocity of a proton accelerated by such a potential? (b) An electron?

66. Suppose you use an average of 500 kW·h of electric energy per month in your home. (a) How long would 1.00 g of mass converted to electric energy with an efficiency of 38.0% last you? (b) How many homes could be supplied at the 500 kW·h per month rate for one year by the energy from the described mass conversion?

67. (a) Using data from Table 7.1, find the mass destroyed when the energy in a barrel of crude oil is released. (b) Given these barrels contain 200 liters and assuming the density of crude oil is $750 \text{ kg/m}^3$, what is the ratio of mass destroyed to original mass, $\Delta m / m$?

68. (a) Calculate the energy released by the destruction of 1.00 kg of mass. (b) How many kilograms could be lifted to a 10.0 km height by this amount of energy?

69. A Van de Graaff accelerator utilizes a 50.0 MV potential difference to accelerate charged particles such as protons. (a) What is the velocity of a proton accelerated by such a potential? (b) An electron?

70. Suppose you use an average of 500 kW·h of electric energy per month in your home. (a) How long would 1.00 g of mass converted to electric energy with an efficiency of 38.0% last you? (b) How many homes could be supplied at the 500 kW·h per month rate for one year by the energy from the described mass conversion?

71. (a) A nuclear power plant converts energy from nuclear fission into electricity with an efficiency of 35.0%. How much mass is destroyed in one year to produce a continuous 1000 MW of electric power? (b) Do you think it would be possible to observe this mass loss if the total mass of the fuel is $10^4 \text{ kg}$?

72. Nuclear-powered rockets were researched for some years before safety concerns became paramount. (a) What fraction of a rocket’s mass would have to be destroyed to get it into a low Earth orbit, neglecting the decrease in gravity? (Assume an orbital altitude of 250 km, and calculate both the kinetic energy (classical) and the gravitational potential energy needed.) (b) If the ship has a mass of $1.00 \times 10^5 \text{ kg}$ (100 tons), what total yield nuclear explosion in tons of TNT is needed?
29 INTRODUCTION TO QUANTUM PHYSICS

Figure 29.1 A black fly imaged by an electron microscope is as monstrous as any science-fiction creature. (credit: U.S. Department of Agriculture via Wikimedia Commons)

Learning Objectives

29.1. Quantization of Energy
- Explain Max Planck’s contribution to the development of quantum mechanics.
- Explain why atomic spectra indicate quantization.

29.2. The Photoelectric Effect
- Describe a typical photoelectric-effect experiment.
- Determine the maximum kinetic energy of photoelectrons ejected by photons of one energy or wavelength, when given the maximum kinetic energy of photoelectrons for a different photon energy or wavelength.

29.3. Photon Energies and the Electromagnetic Spectrum
- Explain the relationship between the energy of a photon in joules or electron volts and its wavelength or frequency.
- Calculate the number of photons per second emitted by a monochromatic source of specific wavelength and power.

29.4. Photon Momentum
- Relate the linear momentum of a photon to its energy or wavelength, and apply linear momentum conservation to simple processes involving the emission, absorption, or reflection of photons.
- Account qualitatively for the increase of photon wavelength that is observed, and explain the significance of the Compton wavelength.

29.5. The Particle-Wave Duality
- Explain what the term particle-wave duality means, and why it is applied to EM radiation.

29.6. The Wave Nature of Matter
- Describe the Davisson-Germer experiment, and explain how it provides evidence for the wave nature of electrons.

29.7. Probability: The Heisenberg Uncertainty Principle
- Use both versions of Heisenberg’s uncertainty principle in calculations.
- Explain the implications of Heisenberg’s uncertainty principle for measurements.

29.8. The Particle-Wave Duality Reviewed
- Explain the concept of particle-wave duality, and its scope.
Introduction to Quantum Physics

Quantum mechanics is the branch of physics needed to deal with submicroscopic objects. Because these objects are smaller than we can observe directly with our senses and generally must be observed with the aid of instruments, parts of quantum mechanics seem as foreign and bizarre as parts of relativity. But, like relativity, quantum mechanics has been shown to be valid—truth is often stranger than fiction.

Certain aspects of quantum mechanics are familiar to us. We accept as fact that matter is composed of atoms, the smallest unit of an element, and that these atoms combine to form molecules, the smallest unit of a compound. (See Figure 29.2.) While we cannot see the individual water molecules in a stream, for example, we are aware that this is because molecules are so small and so numerous in that stream. When introducing atoms, we commonly say that electrons orbit atoms in discrete shells around a tiny nucleus, itself composed of smaller particles called protons and neutrons. We are also aware that electric charge comes in tiny units carried almost entirely by electrons and protons. As with water molecules in a stream, we do not notice individual charges in the current through a light bulb, because the charges are so small and so numerous in the macroscopic situations we sense directly.

Figure 29.2 Atoms and their substructure are familiar examples of objects that require quantum mechanics to be fully explained. Certain of their characteristics, such as the discrete electron shells, are classical physics explanations. In quantum mechanics we conceptualize discrete "electron clouds" around the nucleus.

Making Connections: Realms of Physics

Classical physics is a good approximation of modern physics under conditions first discussed in the The Nature of Science and Physics. Quantum mechanics is valid in general, and it must be used rather than classical physics to describe small objects, such as atoms.

Atoms, molecules, and fundamental electron and proton charges are all examples of physical entities that are quantized—that is, they appear only in certain discrete values and do not have every conceivable value. Quantized is the opposite of continuous. We cannot have a fraction of an atom, or part of an electron's charge, or 14-1/3 cents, for example. Rather, everything is built of integral multiples of these substructures. Quantum physics is the branch of physics that deals with small objects and the quantization of various entities, including energy and angular momentum. Just as with classical physics, quantum physics has several subfields, such as mechanics and the study of electromagnetic forces. The correspondence principle states that in the classical limit (large, slow-moving objects), quantum mechanics becomes the same as classical physics. In this chapter, we begin the development of quantum mechanics and its description of the strange submicroscopic world. In later chapters, we will examine many areas, such as atomic and nuclear physics, in which quantum mechanics is crucial.

29.1 Quantization of Energy

Planck's Contribution

Energy is quantized in some systems, meaning that the system can have only certain energies and not a continuum of energies, unlike the classical case. This would be like having only certain speeds at which a car can travel because its kinetic energy can have only certain values. We also find that some forms of energy transfer take place with discrete lumps of energy. While most of us are familiar with the quantization of matter into lumps called atoms, molecules, and the like, we are less aware that energy, too, can be quantized. Some of the earliest clues about the necessity of quantum mechanics over classical physics came from the quantization of energy.
Figure 29.3 Graphs of blackbody radiation (from an ideal radiator) at three different radiator temperatures. The intensity or rate of radiation emission increases dramatically with temperature, and the peak of the spectrum shifts toward the visible and ultraviolet parts of the spectrum. The shape of the spectrum cannot be described with classical physics.

Where is the quantization of energy observed? Let us begin by considering the emission and absorption of electromagnetic (EM) radiation. The EM spectrum radiated by a hot solid is linked directly to the solid’s temperature. (See Figure 29.3.) An ideal radiator is one that has an emissivity of 1 at all wavelengths and, thus, is jet black. Ideal radiators are therefore called blackbodies, and their EM radiation is called blackbody radiation. It was discussed that the total intensity of the radiation varies as $T^{-4}$, the fourth power of the absolute temperature of the body, and that the peak of the spectrum shifts to shorter wavelengths at higher temperatures. All of this seems quite continuous, but it was the curve of the spectrum of intensity versus wavelength that gave a clue that the energies of the atoms in the solid are quantized. In fact, providing a theoretical explanation for the experimentally measured shape of the spectrum was a mystery at the turn of the century. When this “ultraviolet catastrophe” was eventually solved, the answers led to new technologies such as computers and the sophisticated imaging techniques described in earlier chapters. Once again, physics as an enabling science changed the way we live.

The German physicist Max Planck (1858–1947) used the idea that atoms and molecules in a body act like oscillators to absorb and emit radiation. The energies of the oscillating atoms and molecules had to be quantized to correctly describe the shape of the blackbody spectrum. Planck deduced that the energy of an oscillator having a frequency $f$ is given by

$$E = \left(n + \frac{1}{2}\right)\hbar f.$$  

Here $n$ is any nonnegative integer (0, 1, 2, 3, …). The symbol $\hbar$ stands for Planck’s constant, given by

$$\hbar = 6.626 \times 10^{-34} \text{ J} \cdot \text{s}.$$  

The equation $E = \left(n + \frac{1}{2}\right)\hbar f$ means that an oscillator having a frequency $f$ (emitting and absorbing EM radiation of frequency $f$) can have its energy increase or decrease only in discrete steps of size

$$\Delta E = h f.$$  

It might be helpful to mention some macroscopic analogies of this quantization of energy phenomena. This is like a pendulum that has a characteristic oscillation frequency but can swing with only certain amplitudes. Quantization of energy also resembles a standing wave on a string that allows only particular harmonics described by integers. It is also similar to going up and down a hill using discrete stair steps rather than being able to move up and down a continuous slope. Your potential energy takes on discrete values as you move from step to step.

Using the quantization of oscillators, Planck was able to correctly describe the experimentally known shape of the blackbody spectrum. This was the first indication that energy is sometimes quantized on a small scale and earned him the Nobel Prize in Physics in 1918. Although Planck’s theory comes from observations of a macroscopic object, its analysis is based on atoms and molecules. It was such a revolutionary departure from classical physics that Planck himself was reluctant to accept his own idea that energy states are not continuous. The general acceptance of Planck’s energy quantization was greatly enhanced by Einstein’s explanation of the photoelectric effect (discussed in the next section), which took energy quantization a step further. Planck was fully involved in the development of both early quantum mechanics and relativity. He quickly embraced Einstein’s special relativity, published in 1905, and in 1906 Planck was the first to suggest the correct formula for relativistic momentum, $p = \gamma mu$.
Figure 29.4 The German physicist Max Planck had a major influence on the early development of quantum mechanics, being the first to recognize that energy is sometimes quantized. Planck also made important contributions to special relativity and classical physics. (credit: Library of Congress, Prints and Photographs Division via Wikimedia Commons)

Note that Planck's constant $h$ is a very small number. So for an infrared frequency of $10^{14}$ Hz being emitted by a blackbody, for example, the difference between energy levels is only $\Delta E = hf = (6.63 \times 10^{-34} \text{ J} \cdot \text{s})(10^{14} \text{ Hz}) = 6.63 \times 10^{-20} \text{ J}$, or about 0.4 eV. This 0.4 eV of energy is significant compared with typical atomic energies, which are on the order of an electron volt, or thermal energies, which are typically fractions of an electron volt. But on a macroscopic or classical scale, energies are typically on the order of joules. Even if macroscopic energies are quantized, the quantum steps are too small to be noticed. This is an example of the correspondence principle. For a large object, quantum mechanics produces results indistinguishable from those of classical physics.

Atomic Spectra

Now let us turn our attention to the emission and absorption of EM radiation by gases. The Sun is the most common example of a body containing gases emitting an EM spectrum that includes visible light. We also see examples in neon signs and candle flames. Studies of emissions of hot gases began more than two centuries ago, and it was soon recognized that these emission spectra contained huge amounts of information. The type of gas and its temperature, for example, could be determined. We now know that these EM emissions come from electrons transitioning between energy levels in individual atoms and molecules; thus, they are called atomic spectra. Atomic spectra remain an important analytical tool today. Figure 29.5 shows an example of an emission spectrum obtained by passing an electric discharge through a material. One of the most important characteristics of these spectra is that they are discrete. By this we mean that only certain wavelengths, and hence frequencies, are emitted. This is called a line spectrum. If frequency and energy are associated as $\Delta E = hf$, the energies of the electrons in the emitting atoms and molecules are quantized. This is discussed in more detail later in this chapter.

Figure 29.5 Emission spectrum of oxygen. When an electrical discharge is passed through a substance, its atoms and molecules absorb energy, which is reemitted as EM radiation. The discrete nature of these emissions implies that the energy states of the atoms and molecules are quantized. Such atomic spectra were used as analytical tools for many decades before it was understood why they are quantized. (credit: Teravolt, Wikimedia Commons)

It was a major puzzle that atomic spectra are quantized. Some of the best minds of 19th-century science failed to explain why this might be. Not until the second decade of the 20th century did an answer based on quantum mechanics begin to emerge. Again a macroscopic or classical body of gas was involved in the studies, but the effect, as we shall see, is due to individual atoms and molecules.

PhET Explorations: Models of the Hydrogen Atom

How did scientists figure out the structure of atoms without looking at them? Try out different models by shooting light at the atom. Check how the prediction of the model matches the experimental results.

Figure 29.6 Models of the Hydrogen Atom (http://cnx.org/content/m42554/1.4/hydrogen-atom_en.jar)

29.2 The Photoelectric Effect

When light strikes materials, it can eject electrons from them. This is called the photoelectric effect, meaning that light (photo) produces electricity. One common use of the photoelectric effect is in light meters, such as those that adjust the automatic iris on various types of cameras. In a similar way, another use is in solar cells, as you probably have in your calculator or have seen on a roof top or a roadside sign. These make use of the photoelectric effect to convert light into electricity for running different devices.
This effect has been known for more than a century and can be studied using a device such as that shown in Figure 29.7. This figure shows an evacuated tube with a metal plate and a collector wire that are connected by a variable voltage source, with the collector more negative than the plate. When light (or other EM radiation) strikes the plate in the evacuated tube, it may eject electrons. If the electrons have energy in electron volts (eV) greater than the potential difference between the plate and the wire in volts, some electrons will be collected on the wire. Since the electron energy in eV is \( qV \), where \( q \) is the electron charge and \( V \) is the potential difference, the electron energy can be measured by adjusting the retarding voltage between the wire and the plate. The voltage that stops the electrons from reaching the wire equals the energy in eV. For example, if \(-3.00\) V barely stops the electrons, their energy is 3.00 eV. The number of electrons ejected can be determined by measuring the current between the wire and plate. The more light, the more electrons; a little circuitry allows this device to be used as a light meter.

What is really important about the photoelectric effect is what Albert Einstein deduced from it. Einstein realized that there were several characteristics of the photoelectric effect that could be explained only if EM radiation is itself quantized: the apparently continuous stream of energy in an EM wave is actually composed of energy quanta called photons. In his explanation of the photoelectric effect, Einstein defined a quantized unit or quantum of EM energy, which we now call a photon, with an energy proportional to the frequency of EM radiation. In equation form, the photon energy is

\[
E = hf,
\]

where \( E \) is the energy of a photon of frequency \( f \) and \( h \) is Planck’s constant. This revolutionary idea looks similar to Planck’s quantization of energy states in blackbody oscillators, but it is quite different. It is the quantization of EM radiation itself. EM waves are composed of photons and are not continuous smooth waves as described in previous chapters on optics. Their energy is absorbed and emitted in lumps, not continuously. This is exactly consistent with Planck’s quantization of energy levels in blackbody oscillators, since these oscillators increase and decrease their energy in steps of \( hf \) by absorbing and emitting photons having \( E = hf \). We do not observe this with our eyes, because there are so many photons in common light sources that individual photons go unnoticed. (See Figure 29.8.) The next section of the text (Photon Energies and the Electromagnetic Spectrum) is devoted to a discussion of photons and some of their characteristics and implications. For now, we will use the photon concept to explain the photoelectric effect, much as Einstein did.

The photoelectric effect has the properties discussed below. All these properties are consistent with the idea that individual photons of EM radiation are absorbed by individual electrons in a material, with the electron gaining the photon’s energy. Some of these properties are inconsistent with the idea that EM radiation is a simple wave. For simplicity, let us consider what happens with monochromatic EM radiation in which all photons have the same energy \( hf \).

1. If we vary the frequency of the EM radiation falling on a material, we find the following: For a given material, there is a threshold frequency \( f_0 \) for the EM radiation below which no electrons are ejected, regardless of intensity. Individual photons interact with individual electrons. Thus if the photon energy is too small to break an electron away, no electrons will be ejected. If EM radiation was a simple wave, sufficient energy could be obtained by increasing the intensity.

2. Once EM radiation falls on a material, electrons are ejected without delay. As soon as an individual photon of a sufficiently high frequency is absorbed by an individual electron, the electron is ejected. If the EM radiation were a simple wave, several minutes would be required for sufficient energy to be deposited to the metal surface to eject an electron.

3. The number of electrons ejected per unit time is proportional to the intensity of the EM radiation and to no other characteristic. High-intensity EM radiation consists of large numbers of photons per unit area, with all photons having the same characteristic energy \( hf \).

4. If we vary the intensity of the EM radiation and measure the energy of ejected electrons, we find the following: The maximum kinetic energy of ejected electrons is independent of the intensity of the EM radiation. Since there are so many electrons in a material, it is extremely unlikely that two photons will interact with the same electron at the same time, thereby increasing the energy given it. Instead (as noted in 3 above), increased intensity results in more electrons of the same energy being ejected. If EM radiation were a simple wave, a higher intensity could give more energy, and higher-energy electrons would be ejected.
5. The kinetic energy of an ejected electron equals the photon energy minus the binding energy of the electron in the specific material. An individual photon can give all of its energy to an electron. The photon's energy is partly used to break the electron away from the material. The remainder goes into the ejected electron's kinetic energy. In equation form, this is given by

\[ KE_e = hf - BE, \]  

where \( KE_e \) is the maximum kinetic energy of the ejected electron, \( hf \) is the photon's energy, and \( BE \) is the binding energy of the electron to the particular material. (\( BE \) is sometimes called the work function of the material.) This equation, due to Einstein in 1905, explains the properties of the photoelectric effect quantitatively. An individual photon of EM radiation (it does not come any other way) interacts with an individual electron, supplying enough energy, \( BE \), to break it away, with the remainder going to kinetic energy. The binding energy is \( BE = hf_0 \), where \( f_0 \) is the threshold frequency for the particular material. Figure 29.9 shows a graph of maximum \( KE_e \) versus the frequency of incident EM radiation falling on a particular material.

\[ KE_e = hf - BE \]

\[ f_0 = \frac{BE}{h} \]

\[ E = hf \]  

Since we are given the wavelength rather than the frequency, we solve the familiar relationship \( c = \frac{E}{f} \) for the frequency, yielding

\[ f = \frac{c}{\lambda}. \]  

Combining these two equations gives the useful relationship

\[ E = \frac{hc}{\lambda}. \]  

Now substituting known values yields

\[ E = \left( \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{420 \times 10^{-9} \text{ m}} \right) \left( 3.00 \times 10^8 \text{ m/s} \right) = 4.74 \times 10^{-19} \text{ J}. \]
Converting to eV, the energy of the photon is
\[ E = \left(4.74 \times 10^{-19} \text{ J}\right) \frac{1 \text{ eV}}{1.6 \times 10^{-19} \text{ J}} = 2.96 \text{ eV}. \] (29.10)

**Solution for (b)**
Finding the kinetic energy of the ejected electron is now a simple application of the equation \( KE_e = hf - BE \). Substituting the photon energy and binding energy yields
\[ KE_e = hf - BE = 2.96 \text{ eV} - 2.71 \text{ eV} = 0.246 \text{ eV}. \] (29.11)

**Discussion**
The energy of this 420-nm photon of violet light is a tiny fraction of a joule, and so it is no wonder that a single photon would be difficult for us to sense directly—humans are more attuned to energies on the order of joules. But looking at the energy in electron volts, we can see that this photon has enough energy to affect atoms and molecules. A DNA molecule can be broken with about 1 eV of energy, for example, and typical atomic and molecular energies are on the order of eV, so that the UV photon in this example could have biological effects. The ejected electron (called a photoelectron) has a rather low energy, and it would not travel far, except in a vacuum. The electron would be stopped by a retarding potential of but 0.26 eV. In fact, if the photon wavelength were longer and its energy less than 2.71 eV, then the formula would give a negative kinetic energy, an impossibility. This simply means that the 420-nm photons with their 2.96-eV energy are not much above the frequency threshold. You can show for yourself that the threshold wavelength is 459 nm (blue light). This means that if calcium metal is used in a light meter, the meter will be insensitive to wavelengths longer than those of blue light. Such a light meter would be completely insensitive to red light, for example.

---

PhET Explorations: Photoelectric Effect

See how light knocks electrons off a metal target, and recreate the experiment that spawned the field of quantum mechanics.

PhET Interactive Simulation

Figure 29.10 Photoelectric Effect (http://cnx.org/content/m42558/1.4/photoelectric_en.jar)

### 29.3 Photon Energies and the Electromagnetic Spectrum

**Ionizing Radiation**

A photon is a quantum of EM radiation. Its energy is given by \( E = hf \) and is related to the frequency \( f \) and wavelength \( \lambda \) of the radiation by
\[ E = hf = \frac{hc}{\lambda} \text{(energy of a photon)}, \] (29.12)

where \( E \) is the energy of a single photon and \( c \) is the speed of light. When working with small systems, energy in eV is often useful. Note that Planck’s constant in these units is
\[ h = 4.14 \times 10^{-15} \text{ eV} \cdot \text{s}. \] (29.13)

Since many wavelengths are stated in nanometers (nm), it is also useful to know that
\[ hc = 1240 \text{ eV} \cdot \text{nm}. \] (29.14)

These will make many calculations a little easier.

All EM radiation is composed of photons. Figure 29.11 shows various divisions of the EM spectrum plotted against wavelength, frequency, and photon energy. Previously in this book, photon characteristics were alluded to in the discussion of some of the characteristics of UV, x rays, and \( \gamma \) rays, the first of which start with frequencies just above violet in the visible spectrum. It was noted that these types of EM radiation have characteristics much different than visible light. We can now see that such properties arise because photon energy is larger at high frequencies.
Figure 29.11 The EM spectrum, showing major categories as a function of photon energy in eV, as well as wavelength and frequency. Certain characteristics of EM radiation are directly attributable to photon energy alone.

Table 29.1 Representative Energies for Submicroscopic Effects (Order of Magnitude Only)

<table>
<thead>
<tr>
<th>Energy Description</th>
<th>Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotational energies of molecules</td>
<td>$10^{-5}$</td>
</tr>
<tr>
<td>Vibrational energies of molecules</td>
<td>0.1</td>
</tr>
<tr>
<td>Energy between outer electron shells in atoms</td>
<td>1</td>
</tr>
<tr>
<td>Binding energy of a weakly bound molecule</td>
<td>1</td>
</tr>
<tr>
<td>Energy of red light</td>
<td>2</td>
</tr>
<tr>
<td>Binding energy of a tightly bound molecule</td>
<td>10</td>
</tr>
<tr>
<td>Energy to ionize atom or molecule</td>
<td>10 to 1000</td>
</tr>
</tbody>
</table>

Photons act as individual quanta and interact with individual electrons, atoms, molecules, and so on. The energy a photon carries is, thus, crucial to the effects it has. Table 29.1 lists representative submicroscopic energies in eV. When we compare photon energies from the EM spectrum in Figure 29.11 with energies in the table, we can see how effects vary with the type of EM radiation.

Gamma rays, a form of nuclear and cosmic EM radiation, can have the highest frequencies and, hence, the highest photon energies in the EM spectrum. For example, a $\gamma$-ray photon with $f = 10^{21}$ Hz has an energy $E = hf = 6.63 \times 10^{-13} J = 4.14$ MeV. This is sufficient energy to ionize thousands of atoms and molecules, since only 10 to 1000 eV are needed per ionization. In fact, $\gamma$ rays are one type of ionizing radiation, as are x rays and UV, because they produce ionization in materials that absorb them. Because so much ionization can be produced, a single $\gamma$-ray photon can cause significant damage to biological tissue, killing cells or damaging their ability to properly reproduce. When cell reproduction is disrupted, the result can be cancer, one of the known effects of exposure to ionizing radiation. Since cancer cells are rapidly reproducing, they are exceptionally sensitive to the disruption produced by ionizing radiation. This means that ionizing radiation has positive uses in cancer treatment as well as risks in producing cancer.
High photon energy also enables $\gamma$ rays to penetrate materials, since a collision with a single atom or molecule is unlikely to absorb all the $\gamma$ ray's energy. This can make $\gamma$ rays useful as a probe, and they are sometimes used in medical imaging. x rays, as you can see in Figure 29.11, overlap with the low-frequency end of the $\gamma$ ray range. Since x rays have energies of keV and up, individual x-ray photons also can produce large amounts of ionization. At lower photon energies, x rays are not as penetrating as $\gamma$ rays and are slightly less hazardous. X rays are ideal for medical imaging, their most common use, and a fact that was recognized immediately upon their discovery in 1895 by the German physicist W. C. Roentgen (1845–1923). (See Figure 29.12.) Within one year of their discovery, x rays (for a time called Roentgen rays) were used for medical diagnostics. Roentgen received the 1901 Nobel Prize for the discovery of x rays.

Connections: Conservation of Energy

Once again, we find that conservation of energy allows us to consider the initial and final forms that energy takes, without having to make detailed calculations of the intermediate steps. Example 29.2 is solved by considering only the initial and final forms of energy.

Figure 29.13 X rays are produced when energetic electrons strike the copper anode of this cathode ray tube (CRT). Electrons (shown here as separate particles) interact individually with the material they strike, sometimes producing photons of EM radiation.
While γ rays originate in nuclear decay, x rays are produced by the process shown in Figure 29.13. Electrons ejected by thermal agitation from a hot filament in a vacuum tube are accelerated through a high voltage, gaining kinetic energy from the electrical potential energy. When they strike the anode, the electrons convert their kinetic energy to a variety of forms, including thermal energy. But since an accelerated charge radiates EM waves, and since the electrons act individually, photons are also produced. Some of these x-ray photons obtain the kinetic energy of the electron. The accelerated electrons originate at the cathode, so such a tube is called a cathode ray tube (CRT), and various versions of them are found in older TV and computer screens as well as in x-ray machines.

Example 29.2 X-ray Photon Energy and X-ray Tube Voltage

Find the maximum energy in eV of an x-ray photon produced by electrons accelerated through a potential difference of 50.0 kV in a CRT like the one in Figure 29.13.

Strategy
Electrons can give all of their kinetic energy to a single photon when they strike the anode of a CRT. (This is something like the photoelectric effect in reverse.) The kinetic energy of the electron comes from electrical potential energy. Thus we can simply equate the maximum photon energy to the electrical potential energy—that is, \( hf = qV \). (We do not have to calculate each step from beginning to end if we know that all of the starting energy \( qV \) is converted to the final form \( hf \).)

Solution
The maximum photon energy is \( hf = qV \), where \( q \) is the charge of the electron and \( V \) is the accelerating voltage. Thus,

\[
hf = (1.60 \times 10^{-19} \text{C})(50.0 \times 10^3 \text{V}).
\]  

From the definition of the electron volt, we know \( 1 \text{ eV} = 1.60 \times 10^{-19} \text{J} \), where \( 1 \text{J} = 1 \text{C} \cdot \text{V} \). Gathering factors and converting energy to eV yields

\[
hf = (50.0 \times 10^3)(1.60 \times 10^{-19} \text{C} \cdot \text{V}) = (50.0 \times 10^3)(1 \text{ eV}) = 50.0 \text{ keV}.
\]

Discussion
This example produces a result that can be applied to many similar situations. If you accelerate a single elementary charge, like that of an electron, through a potential given in volts, then its energy in eV has the same numerical value. Thus a 50.0-kV potential generates 50.0 keV electrons, which in turn can produce photons with a maximum energy of 50 keV. Similarly, a 100-kV potential in an x-ray tube can generate up to 100-keV x-ray photons. Many x-ray tubes have adjustable voltages so that various energy x rays with differing energies, and therefore differing abilities to penetrate, can be generated.

![X-ray spectrum](image)

**Figure 29.14** X-ray spectrum obtained when energetic electrons strike a material. The smooth part of the spectrum is bremsstrahlung, while the peaks are characteristic of the anode material. Both are atomic processes that produce energetic photons known as x-ray photons.

**Figure 29.14** shows the spectrum of x rays obtained from an x-ray tube. There are two distinct features to the spectrum. First, the smooth distribution results from electrons being decelerated in the anode material. A curve like this is obtained by detecting many photons, and it is apparent that the maximum energy is unlikely. This decelerating process produces radiation that is called bremsstrahlung (German for braking radiation). The second feature is the existence of sharp peaks in the spectrum; these are called characteristic x rays, since they are characteristic of the anode material. Characteristic x rays come from atomic excitations unique to a given type of anode material. They are akin to lines in atomic spectra, implying the energy levels of atoms are quantized. Phenomena such as discrete atomic spectra and characteristic x rays are explored further in Atomic Physics.

**Ultraviolet radiation** (approximately 4 eV to 300 eV) overlaps with the low end of the energy range of x rays, but UV is typically lower in energy. UV comes from the de-excitation of atoms that may be part of a hot solid or gas. These atoms can be given energy that they later release as UV by numerous processes, including electric discharge, nuclear explosion, thermal agitation, and exposure to x rays. A UV photon has sufficient energy to ionize atoms and molecules, which makes its effects different from those of visible light. UV thus has some of the same biological effects as γ rays and x rays. For example, it can cause skin cancer and is used as a sterilizer. The major difference is that several UV photons are required to disrupt cell reproduction or kill a bacterium, whereas single γ-ray and X-ray photons can do the same damage. But since UV does have the energy to alter...
molecules, it can do what visible light cannot. One of the beneficial aspects of UV is that it triggers the production of vitamin D in the skin, whereas visible light has insufficient energy per photon to alter the molecules that trigger this production. Infantile jaundice is treated by exposing the baby to UV (with eye protection), called phototherapy, the beneficial effects of which are thought to be related to its ability to help prevent the buildup of potentially toxic bilirubin in the blood.

**Example 29.3 Photon Energy and Effects for UV**

Short-wavelength UV is sometimes called vacuum UV, because it is strongly absorbed by air and must be studied in a vacuum. Calculate the photon energy in eV for 100-nm vacuum UV, and estimate the number of molecules it could ionize or break apart.

**Strategy**

Using the equation $E = hf$ and appropriate constants, we can find the photon energy and compare it with energy information in Table 29.1.

**Solution**

The energy of a photon is given by

$$E = hf = \frac{hc}{\lambda}.$$  

Using $hc = 1240 \text{ eV} \cdot \text{nm}$, we find that

$$E = \frac{hc}{\lambda} = \frac{1240 \text{ eV} \cdot \text{nm}}{100 \text{ nm}} = 12.4 \text{ eV}.$$  

**Discussion**

According to Table 29.1, this photon energy might be able to ionize an atom or molecule, and it is about what is needed to break up a tightly bound molecule, since they are bound by approximately 10 eV. This photon energy could destroy about a dozen weakly bound molecules. Because of its high photon energy, UV disrupts atoms and molecules it interacts with. One good consequence is that all but the longest-wavelength UV is strongly absorbed and is easily blocked by sunglasses. In fact, most of the Sun's UV is absorbed by a thin layer of ozone in the upper atmosphere, protecting sensitive organisms on Earth. Damage to our ozone layer by the addition of such chemicals as CFC's has reduced this protection for us.

**Visible Light**

The range of photon energies for *visible light* from red to violet is 1.63 to 3.26 eV, respectively (left for this chapter’s Problems and Exercises to verify). These energies are on the order of those between outer electron shells in atoms and molecules. This means that these photons can be absorbed by atoms and molecules. A single photon can actually stimulate the retina, for example, by altering a receptor molecule that then triggers a nerve impulse. Photons can be absorbed or emitted only by atoms and molecules that have precisely the correct quantized energy step to do so. For example, if a red photon of frequency $f$ encounters a molecule that has an energy step, $\Delta E$, equal to $hf$, then the photon can be absorbed. Violet flowers absorb red and reflect violet; this implies there is no energy step between levels in the receptor molecule equal to the violet photon’s energy, but there is an energy step for the red.

There are some noticeable differences in the characteristics of light between the two ends of the visible spectrum that are due to photon energies. Red light has insufficient photon energy to expose most black-and-white film, and it is thus used to illuminate darkrooms where such film is developed. Since violet light has a higher photon energy, dyes that absorb violet tend to fade more quickly than those that do not. (See Figure 29.15.) Take a look at some faded color posters in a storefront some time, and you will notice that the blues and violets are the last to fade. This is because other dyes, such as red and green dyes, absorb blue and violet photons, the higher energies of which break up their weakly bound molecules. (Complex molecules such as those in dyes and DNA tend to be weakly bound.) Blue and violet dyes reflect those colors and, therefore, do not absorb these more energetic photons, thus suffering less molecular damage.
Transparent materials, such as some glasses, do not absorb any visible light, because there is no energy step in the atoms or molecules that could absorb the light. Since individual photons interact with individual atoms, it is nearly impossible to have two photons absorbed simultaneously to reach a large energy step. Because of its lower photon energy, visible light can sometimes pass through many kilometers of a substance, while higher frequencies like UV, x-ray, and γ-rays are absorbed, because they have sufficient photon energy to ionize the material.

**Example 29.4 How Many Photons per Second Does a Typical Light Bulb Produce?**

Assuming that 10.0% of a 100-W light bulb's energy output is in the visible range (typical for incandescent bulbs) with an average wavelength of 580 nm, calculate the number of visible photons emitted per second.

**Strategy**

Power is energy per unit time, and so if we can find the energy per photon, we can determine the number of photons per second. This will best be done in joules, since power is given in watts, which are joules per second.

**Solution**

The power in visible light production is 10.0% of 100 W, or 10.0 J/s. The energy of the average visible photon is found by substituting the given average wavelength into the formula

\[ E = \frac{hc}{\lambda}. \]  

This produces

\[ E = \frac{(6.63 \times 10^{-34} \text{ J} \cdot \text{s})(3.00 \times 10^8 \text{ m/s})}{580 \times 10^{-9} \text{ m}} = 3.43 \times 10^{-19} \text{ J}. \]

The number of visible photons per second is thus

\[ \text{photon/s} = \frac{10.0 \text{ J/s}}{3.43 \times 10^{-19} \text{ J/phot on}} = 2.92 \times 10^{19} \text{ photon/s}. \]

**Discussion**

This incredible number of photons per second is verification that individual photons are insignificant in ordinary human experience. It is also a verification of the correspondence principle—on the macroscopic scale, quantization becomes essentially continuous or classical. Finally, there are so many photons emitted by a 100-W light bulb that it can be seen by the unaided eye many kilometers away.

**Lower-Energy Photons**

**Infrared radiation (IR)** has even lower photon energies than visible light and cannot significantly alter atoms and molecules. IR can be absorbed and emitted by atoms and molecules, particularly between closely spaced states. IR is extremely strongly absorbed by water, for example, because water molecules have many states separated by energies on the order of $10^{-5}$ eV to $10^{-2}$ eV, well within the IR and microwave energy ranges. This is why in the IR range, skin is almost jet black, with an emissivity near 1—there are many states in water molecules in the skin that can absorb a large range of IR photon energies. Not all molecules have this property. Air, for example, is nearly transparent to many IR frequencies.
Microwaves are the highest frequencies that can be produced by electronic circuits, although they are also produced naturally. Thus microwaves are similar to IR but do not extend to as high frequencies. There are states in water and other molecules that have the same frequency and energy as microwaves, typically about $10^{-5}$ eV. This is one reason why food absorbs microwaves more strongly than many other materials, making microwave ovens an efficient way of putting energy directly into food.

Photon energies for both IR and microwaves are so low that huge numbers of photons are involved in any significant energy transfer by IR or microwaves (such as warming yourself with a heat lamp or cooking pizza in the microwave). Visible light, IR, microwaves, and all lower frequencies cannot produce ionization with single photons and do not ordinarily have the hazards of higher frequencies. When visible, IR, or microwave radiation is hazardous, such as the inducement of cataracts by microwaves, the hazard is due to huge numbers of photons acting together (not to an accumulation of photons, such as sterilization by weak UV). The negative effects of visible, IR, or microwave radiation can be thermal effects, which could be produced by any heat source. But one difference is that at very high intensity, strong electric and magnetic fields can be produced by photons acting together. Such electromagnetic fields (EMF) can actually ionize materials.

Misconception Alert: High-Voltage Power Lines

Although some people think that living near high-voltage power lines is hazardous to one’s health, ongoing studies of the transient field effects produced by these lines show their strengths to be insufficient to cause damage. Demographic studies also fail to show significant correlation of ill effects with high-voltage power lines. The American Physical Society issued a report over 10 years ago on power-line fields, which concluded that the scientific literature and reviews of panels show no consistent, significant link between cancer and power-line fields. They also felt that the “diversion of resources to eliminate a threat which has no persuasive scientific basis is disturbing.”

It is virtually impossible to detect individual photons having frequencies below microwave frequencies, because of their low photon energy. But the photons are there. A continuous EM wave can be modeled as photons. At low frequencies, EM waves are generally treated as time- and position-varying electric and magnetic fields with no discernible quantization. This is another example of the correspondence principle in situations involving huge numbers of photons.

PhET Explorations: Color Vision

Make a whole rainbow by mixing red, green, and blue light. Change the wavelength of a monochromatic beam or filter white light. View the light as a solid beam, or see the individual photons.

PhET Interactive Simulation

Figure 29.16 Color Vision (http://cnx.org/content/m42563/1.5/color-vision_en.jar)

29.4 Photon Momentum

Measuring Photon Momentum

The quantum of EM radiation we call a photon has properties analogous to those of particles we can see, such as grains of sand. A photon interacts as a unit in collisions or when absorbed, rather than as an extensive wave. Massive quanta, like electrons, also act like macroscopic particles—something we expect, because they are the smallest units of matter. Particles carry momentum as well as energy. Despite photons having no mass, there has long been evidence that EM radiation carries momentum. (Maxwell and others who studied EM waves predicted that they would carry momentum.) It is now a well-established fact that photons do have momentum. In fact, photon momentum is suggested by the photoelectric effect, where photons knock electrons out of a substance. Figure 29.17 shows macroscopic evidence of photon momentum.
Figure 29.17 shows a comet with two prominent tails. What most people do not know about the tails is that they always point away from the Sun rather than trailing behind the comet (like the tail of Bo Peep's sheep). Comet tails are composed of gases and dust evaporated from the body of the comet and ionized gas. The dust particles recoil away from the Sun when photons scatter from them. Evidently, photons carry momentum in the direction of their motion (away from the Sun), and some of this momentum is transferred to dust particles in collisions. Gas atoms and molecules in the blue tail are most affected by other particles of radiation, such as protons and electrons emanating from the Sun, rather than by the momentum of photons.

Connections: Conservation of Momentum

Not only is momentum conserved in all realms of physics, but all types of particles are found to have momentum. We expect particles with mass to have momentum, but now we see that massless particles including photons also carry momentum.

Momentum is conserved in quantum mechanics just as it is in relativity and classical physics. Some of the earliest direct experimental evidence of this came from scattering of x-ray photons by electrons in substances, named Compton scattering after the American physicist Arthur H. Compton (1892–1962). Around 1923, Compton observed that x rays scattered from materials had a decreased energy and correctly analyzed this as being due to the scattering of photons from electrons. This phenomenon could be handled as a collision between two particles—a photon and an electron at rest in the material. Energy and momentum are conserved in the collision. (See Figure 29.18) He won a Nobel Prize in 1929 for the discovery of this scattering, now called the **Compton effect**, because it helped prove that **photon momentum** is given by

\[ p = \frac{h}{\lambda}, \]

where \( h \) is Planck's constant and \( \lambda \) is the photon wavelength. (Note that relativistic momentum given as \( p = \gamma mu \) is valid only for particles having mass.)
We can see that photon momentum is small, since $p = \frac{h}{\lambda}$ and $h$ is very small. It is for this reason that we do not ordinarily observe photon momentum. Our mirrors do not recoil when light reflects from them (except perhaps in cartoons). Compton saw the effects of photon momentum because he was observing x rays, which have a small wavelength and a relatively large momentum, interacting with the lightest of particles, the electron.

### Example 29.5 Electron and Photon Momentum Compared

(a) Calculate the momentum of a visible photon that has a wavelength of 500 nm. (b) Find the velocity of an electron having the same momentum. (c) What is the energy of the electron, and how does it compare with the energy of the photon?

#### Strategy

Finding the photon momentum is a straightforward application of its definition: $p = \frac{h}{\lambda}$. If we find the photon momentum is small, then we can assume that an electron with the same momentum will be nonrelativistic, making it easy to find its velocity and kinetic energy from the classical formulas.

#### Solution for (a)

Photon momentum is given by the equation:

$$ p = \frac{h}{\lambda}. \quad (29.23) $$

Entering the given photon wavelength yields

$$ p = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{500 \times 10^{-9} \text{ m}} = 1.33 \times 10^{-27} \text{ kg} \cdot \text{m/s}. \quad (29.24) $$

#### Solution for (b)

Since this momentum is indeed small, we will use the classical expression $p = mv$ to find the velocity of an electron with this momentum. Solving for $v$ and using the known value for the mass of an electron gives

$$ v = \frac{p}{m} = \frac{1.33 \times 10^{-27} \text{ kg} \cdot \text{m/s}}{9.11 \times 10^{-31} \text{ kg}} = 1460 \text{ m/s} \approx 1460 \text{ m/s}. \quad (29.25) $$

#### Solution for (c)

The electron has kinetic energy, which is classically given by

$$ KE_e = \frac{1}{2}mv^2. \quad (29.26) $$

Thus,

$$ KE_e = \frac{1}{2}(9.11 \times 10^{-3} \text{ kg})(1455 \text{ m/s})^2 = 9.64 \times 10^{-25} \text{ J}. \quad (29.27) $$

Converting this to eV by multiplying by $(1 \text{ eV})/(1.602 \times 10^{-19} \text{ J})$ yields

$$ KE_e = 6.02 \times 10^{-6} \text{ eV}. \quad (29.28) $$

The photon energy $E$ is

$$ E = \frac{hc}{\lambda} = \frac{1240 \text{ eV} \cdot \text{nm}}{500 \text{ nm}} = 2.48 \text{ eV}, \quad (29.29) $$
which is about five orders of magnitude greater.

**Discussion**

Photon momentum is indeed small. Even if we have huge numbers of them, the total momentum they carry is small. An electron with the same momentum has a 1460 m/s velocity, which is clearly nonrelativistic. A more massive particle with the same momentum would have an even smaller velocity. This is borne out by the fact that it takes far less energy to give an electron the same momentum as a photon. But on a quantum-mechanical scale, especially for high-energy photons interacting with small masses, photon momentum is significant. Even on a large scale, photon momentum can have an effect if there are enough of them and if there is nothing to prevent the slow recoil of matter. Comet tails are one example, but there are also proposals to build space sails that use huge low-mass mirrors (made of aluminized Mylar) to reflect sunlight. In the vacuum of space, the mirrors would gradually recoil and could actually take spacecraft from place to place in the solar system. (See **Figure 29.19**.)

![Space sail diagram](image)

**Figure 29.19** (a) Space sails have been proposed that use the momentum of sunlight reflecting from gigantic low-mass sails to propel spacecraft about the solar system. A Russian test model of this (the Cosmos 1) was launched in 2005, but did not make it into orbit due to a rocket failure. (b) A U.S. version of this, labeled LightSail-1, is scheduled for trial launches in the first part of this decade. It will have a 40-m² sail. (credit: Kim Newton/NASA)

**Relativistic Photon Momentum**

There is a relationship between photon momentum \( p \) and photon energy \( E \) that is consistent with the relation given previously for the relativistic total energy of a particle as \( E^2 = (pc)^2 + (mc)^2 \). We know \( m \) is zero for a photon, but \( p \) is not, so that \( E^2 = (pc)^2 + (mc)^2 \) becomes

\[
E = pc,
\]

(29.30)

or

\[
p = \frac{E}{c} (\text{photons}).
\]

(29.31)

To check the validity of this relation, note that \( E = hc/\lambda \) for a photon. Substituting this into \( p = E/c \) yields

\[
p = \left( \frac{hc}{\lambda} \right) / c = \frac{h}{\lambda},
\]

(29.32)

as determined experimentally and discussed above. Thus, \( p = E/c \) is equivalent to Compton’s result \( p = h/\lambda \). For a further verification of the relationship between photon energy and momentum, see Example 29.6.

**Photon Detectors**

Almost all detection systems talked about thus far—eyes, photographic plates, photomultiplier tubes in microscopes, and CCD cameras—rely on particle-like properties of photons interacting with a sensitive area. A change is caused and either the change is cascaded or zillions of points are recorded to form an image we detect. These detectors are used in biomedical imaging systems, and there is ongoing research into improving the efficiency of receiving photons, particularly by cooling detection systems and reducing thermal effects.

**Example 29.6 Photon Energy and Momentum**

Show that \( p = E/c \) for the photon considered in the Example 29.5.

**Strategy**

We will take the energy \( E \) found in Example 29.5, divide it by the speed of light, and see if the same momentum is obtained as before.

**Solution**

Given that the energy of the photon is 2.48 eV and converting this to joules, we get

\[
p = \frac{E}{c} = \frac{(2.48 \text{ eV})(1.60 \times 10^{-19} \text{ J/eV})}{3.00 \times 10^8 \text{ m/s}} = 1.33 \times 10^{-27} \text{ kg} \cdot \text{m/s}.
\]

(29.33)
Discussion
This value for momentum is the same as found before (note that unrounded values are used in all calculations to avoid even small rounding errors), an expected verification of the relationship $p = E/c$. This also means the relationship between energy, momentum, and mass given by $E^2 = (pc)^2 + (mc^2)^2$ applies to both matter and photons. Once again, note that $p$ is not zero, even when $m$ is.

Problem-Solving Suggestion
Note that the forms of the constants $h = 4.14 \times 10^{-15}$ eV ⋅ s and $hc = 1240$ eV ⋅ nm may be particularly useful for this section's Problems and Exercises.

29.5 The Particle-Wave Duality
We have long known that EM radiation is a wave, capable of interference and diffraction. We now see that light can be modeled as photons, which are massless particles. This may seem contradictory, since we ordinarily deal with large objects that never act like both wave and particle. An ocean wave, for example, looks nothing like a rock. To understand small-scale phenomena, we make analogies with the large-scale phenomena we observe directly. When we say something behaves like a wave, we mean it shows interference effects analogous to those seen in overlapping water waves. (See Figure 29.20.) Two examples of waves are sound and EM radiation. When we say something behaves like a particle, we mean that it interacts as a discrete unit with no interference effects. Examples of particles include electrons, atoms, and photons of EM radiation. How do we talk about a phenomenon that acts like both a particle and a wave?

![Figure 29.20](http://cnx.org/content/m42573/1.3/quantum-wave-interference_en.jar)

Figure 29.20 (a) The interference pattern for light through a double slit is a wave property understood by analogy to water waves. (b) The properties of photons having quantized energy and momentum and acting as a concentrated unit are understood by analogy to macroscopic particles.

There is no doubt that EM radiation interferes and has the properties of wavelength and frequency. There is also no doubt that it behaves as particles—photons with discrete energy. We call this twofold nature the **particle-wave duality**, meaning that EM radiation has both particle and wave properties. This so-called duality is simply a term for properties of the photon analogous to phenomena we can observe directly, on a macroscopic scale. If this term seems strange, it is because we do not ordinarily observe details on the quantum level directly, and our observations yield either particle or wavelike properties, but never both simultaneously.

Since we have a particle-wave duality for photons, and since we have seen connections between photons and matter in that both have momentum, it is reasonable to ask whether there is a particle-wave duality for matter as well. If the EM radiation we once thought to be a pure wave has particle properties, is it possible that matter has wave properties? The answer is yes. The consequences are tremendous, as we will begin to see in the next section.

**PhET Explorations: Quantum Wave Interference**

When do photons, electrons, and atoms behave like particles and when do they behave like waves? Watch waves spread out and interfere as they pass through a double slit, then get detected on a screen as tiny dots. Use quantum detectors to explore how measurements change the waves and the patterns they produce on the screen.

![PhET Interactive Simulation](http://cnx.org/content/m42573/1.3/quantum-wave-interference_en.jar)
29.6 The Wave Nature of Matter

De Broglie Wavelength

In 1923 a French physics graduate student named Prince Louis-Victor de Broglie (1892–1987) made a radical proposal based on the hope that nature is symmetric. If EM radiation has both particle and wave properties, then nature would be symmetric if matter also had both particle and wave properties. If what we once thought of as an unequivocal wave (EM radiation) is also a particle, then what we think of as an unequivocal particle (matter) may also be a wave. De Broglie’s suggestion, made as part of his doctoral thesis, was so radical that it was greeted with some skepticism. A copy of his thesis was sent to Einstein, who said it was not only probably correct, but that it might be of fundamental importance. With the support of Einstein and a few other prominent physicists, de Broglie was awarded his doctorate.

De Broglie took both relativity and quantum mechanics into account to develop the proposal that all particles have a wavelength, given by

\[ \lambda = \frac{h}{p} \] (matter and photons),

where \( h \) is Planck’s constant and \( p \) is momentum. This is defined to be the de Broglie wavelength. (Note that we already have this for photons, from the equation \( p = \frac{h}{\lambda} \).) The hallmark of a wave is interference. If matter is a wave, then it must exhibit constructive and destructive interference. Why isn’t this ordinarily observed? The answer is that in order to see significant interference effects, a wave must interact with an object about the same size as its wavelength. Since \( h \) is very small, \( \lambda \) is also small, especially for macroscopic objects. A 3-kg bowling ball moving at 10 m/s, for example, has

\[ \lambda = \frac{h}{p} = \frac{(6.63 \times 10^{-34} \text{ J}\cdot\text{s})}{(3 \text{ kg})(10 \text{ m/s})} = 2 \times 10^{-35} \text{ m}. \]

This means that to see its wave characteristics, the bowling ball would have to interact with something about \( 10^{-35} \text{ m} \) in size—far smaller than anything known. When waves interact with objects much larger than their wavelength, they show negligible interference effects and move in straight lines (such as light rays in geometric optics). To get easily observed interference effects from particles of matter, the longest wavelength and hence smallest mass possible would be useful. Therefore, this effect was first observed with electrons.

American physicists Clinton J. Davisson and Lester H. Germer in 1925 and, independently, British physicist G. P. Thomson (son of J. J. Thomson, discoverer of the electron) in 1926 scattered electrons from crystals and found diffraction patterns. These patterns are exactly consistent with interference of electrons having the de Broglie wavelength and are somewhat analogous to light interacting with a diffraction grating. (See Figure 29.22.)

Connections: Waves

All microscopic particles, whether massless, like photons, or having mass, like electrons, have wave properties. The relationship between momentum and wavelength is fundamental for all particles.

De Broglie’s proposal of a wave nature for all particles initiated a remarkably productive era in which the foundations for quantum mechanics were laid. In 1926, the Austrian physicist Erwin Schrödinger (1887–1961) published four papers in which the wave nature of particles was treated explicitly with wave equations. At the same time, many others began important work. Among them was German physicist Werner Heisenberg (1901–1976) who, among many other contributions to quantum mechanics, formulated a mathematical treatment of the wave nature of matter that used matrices rather than wave equations. We will deal with some specifics in later sections, but it is worth noting that de Broglie’s work was a watershed for the development of quantum mechanics. De Broglie was awarded the Nobel Prize in 1929 for his vision, as were Davisson and G. P. Thomson in 1937 for their experimental verification of de Broglie’s hypothesis.

Figure 29.22 This diffraction pattern was obtained for electrons diffracted by crystalline silicon. Bright regions are those of constructive interference, while dark regions are those of destructive interference. (credit: Ndthe, Wikimedia Commons)
Example 29.7 Electron Wavelength versus Velocity and Energy

For an electron having a de Broglie wavelength of 0.167 nm (appropriate for interacting with crystal lattice structures that are about this size): (a) Calculate the electron’s velocity, assuming it is nonrelativistic. (b) Calculate the electron’s kinetic energy in eV.

Strategy

For part (a), since the de Broglie wavelength is given, the electron’s velocity can be obtained from \( \lambda = \frac{h}{p} \) by using the nonrelativistic formula for momentum, \( p = mv \). For part (b), once \( v \) is obtained (and it has been verified that \( v \) is nonrelativistic), the classical kinetic energy is simply \( \frac{1}{2}mv^2 \).

Solution for (a)

Substituting the nonrelativistic formula for momentum \( (p = mv) \) into the de Broglie wavelength gives

\[ \lambda = \frac{h}{p} = \frac{h}{mv}. \]  

Solving for \( v \) gives

\[ v = \frac{h}{m\lambda}. \]

Substituting known values yields

\[ v = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{(9.11 \times 10^{-31} \text{ kg})(0.167 \times 10^{-9} \text{ m})} = 4.36 \times 10^6 \text{ m/s}. \]

Solution for (b)

While fast compared with a car, this electron’s speed is not highly relativistic, and so we can comfortably use the classical formula to find the electron’s kinetic energy and convert it to eV as requested.

\[ \text{KE} = \frac{1}{2}mv^2 \]

\[ = \frac{1}{2}(9.11 \times 10^{-31} \text{ kg})(4.36 \times 10^6 \text{ m/s})^2 \]

\[ = (86.4 \times 10^{-18} \text{ J})(\frac{1 \text{ eV}}{1.602 \times 10^{-19} \text{ J}}) \]

\[ = 54.0 \text{ eV} \]

Discussion

This low energy means that these 0.167-nm electrons could be obtained by accelerating them through a 54-V electrostatic potential, an easy task. The results also confirm the assumption that the electrons are nonrelativistic, since their velocity is just over 1% of the speed of light and the kinetic energy is about 0.01% of the rest energy of an electron (0.511 MeV). If the electrons had turned out to be relativistic, we would have had to use more involved calculations employing relativistic formulas.

Electron Microscopes

One consequence or use of the wave nature of matter is found in the electron microscope. As we have discussed, there is a limit to the detail observed with any probe having a wavelength. Resolution, or observable detail, is limited to about one wavelength. Since a potential of only 54 V can produce electrons with sub-nanometer wavelengths, it is easy to get electrons with much smaller wavelengths than those of visible light (hundreds of nanometers). Electron microscopes can, thus, be constructed to detect much smaller details than optical microscopes. (See Figure 29.23.)

There are basically two types of electron microscopes. The transmission electron microscope (TEM) accelerates electrons that are emitted from a hot filament (the cathode). The beam is broadened and then passes through the sample. A magnetic lens focuses the beam image onto a fluorescent screen, a photographic plate, or (most probably) a CCD (light sensitive camera), from which it is transferred to a computer. The TEM is similar to the optical microscope, but it requires a thin sample examined in a vacuum. However it can resolve details as small as 0.1 nm (10^{-10} m), providing magnifications of 100 million times the size of the original object. The TEM has allowed us to see individual atoms and structure of cell nuclei.

The scanning electron microscope (SEM) provides images by using secondary electrons produced by the primary beam interacting with the surface of the sample (see Figure 29.23). The SEM also uses magnetic lenses to focus the beam onto the sample. However, it moves the beam around electrically to “scan” the sample in the x and y directions. A CCD detector is used to process the data for each electron position, producing images like the one at the beginning of this chapter. The SEM has the advantage of not requiring a thin sample and of providing a 3-D view. However, its resolution is about ten times less than a TEM.
Electrons were the first particles with mass to be directly confirmed to have the wavelength proposed by de Broglie. Subsequently, protons, helium nuclei, neutrons, and many others have been observed to exhibit interference when they interact with objects having sizes similar to their de Broglie wavelength. The de Broglie wavelength for massless particles was well established in the 1920s for photons, and it has since been observed that all massless particles have a de Broglie wavelength $\lambda = h/p$. The wave nature of all particles is a universal characteristic of nature. We shall see in following sections that implications of the de Broglie wavelength include the quantization of energy in atoms and molecules, and an alteration of our basic view of nature on the microscopic scale. The next section, for example, shows that there are limits to the precision with which we may make predictions, regardless of how hard we try. There are even limits to the precision with which we may measure an object’s location or energy.

Making Connections: A Submicroscopic Diffraction Grating

The wave nature of matter allows it to exhibit all the characteristics of other, more familiar, waves. Diffraction gratings, for example, produce diffraction patterns for light that depend on grating spacing and the wavelength of the light. This effect, as with most wave phenomena, is most pronounced when the wave interacts with objects having a size similar to its wavelength. For gratings, this is the spacing between multiple slits.) When electrons interact with a system having a spacing similar to the electron wavelength, they show the same types of interference patterns as light does for diffraction gratings, as shown at top left in Figure 29.24.

Atoms are spaced at regular intervals in a crystal as parallel planes, as shown in the bottom part of Figure 29.24. The spacings between these planes act like the openings in a diffraction grating. At certain incident angles, the paths of electrons scattering from successive planes differ by one wavelength and, thus, interfere constructively. At other angles, the path length differences are not an integral wavelength, and there is partial to total destructive interference. This type of scattering from a large crystal with well-defined lattice planes can produce dramatic interference patterns. It is called Bragg reflection, for the father-and-son team who first explored and analyzed it in some detail. The expanded view also shows the path-length differences and indicates how these depend on incident angle $\theta$ in a manner similar to the diffraction patterns for $x$ rays reflecting from a crystal.
Let us take the spacing between parallel planes of atoms in the crystal to be \( d \). As mentioned, if the path length difference (PLD) for the electrons is a whole number of wavelengths, there will be constructive interference—that is, \( \text{PLD} = n\lambda (n = 1, 2, 3, \ldots) \). Because \( AB = BC = d \sin \theta \), we have constructive interference when \( n\lambda = 2d \sin \theta \). This relationship is called the Bragg equation and applies not only to electrons but also to x rays.

The wavelength of matter is a submicroscopic characteristic that explains a macroscopic phenomenon such as Bragg reflection. Similarly, the wavelength of light is a submicroscopic characteristic that explains the macroscopic phenomenon of diffraction patterns.

### 29.7 Probability: The Heisenberg Uncertainty Principle

#### Probability Distribution

Matter and photons are waves, implying they are spread out over some distance. What is the position of a particle, such as an electron? Is it at the center of the wave? The answer lies in how you measure the position of an electron. Experiments show that you will find the electron at some definite location, unlike a wave. But if you set up exactly the same situation and measure it again, you will find the electron in a different location, often far outside any experimental uncertainty in your measurement. Repeated measurements will display a statistical distribution of locations that appears wavelike. (See Figure 29.25.)
Figure 29.25 The building up of the diffraction pattern of electrons scattered from a crystal surface. Each electron arrives at a definite location, which cannot be precisely predicted. The overall distribution shown at the bottom can be predicted as the diffraction of waves having the de Broglie wavelength of the electrons.

Figure 29.26 Double-slit interference for electrons (a) and photons (b) is identical for equal wavelengths and equal slit separations. Both patterns are probability distributions in the sense that they are built up by individual particles traversing the apparatus, the paths of which are not individually predictable.

After de Broglie proposed the wave nature of matter, many physicists, including Schrödinger and Heisenberg, explored the consequences. The idea quickly emerged that, because of its wave character, a particle’s trajectory and destination cannot be precisely predicted for each particle individually. However, each particle goes to a definite place (as illustrated in Figure 29.25). After compiling enough data, you get a distribution related to the particle’s wavelength and diffraction pattern. There is a certain probability of finding the particle at a given location, and the overall pattern is called a probability distribution. Those who developed quantum mechanics devised equations that predicted the probability distribution in various circumstances.

It is somewhat disquieting to think that you cannot predict exactly where an individual particle will go, or even follow it to its destination. Let us explore what happens if we try to follow a particle. Consider the double-slit patterns obtained for electrons and photons in Figure 29.26. First, we note that these patterns are identical, following \( d \sin \theta = m\lambda \), the equation for double-slit constructive interference developed in Photon Energies and the Electromagnetic Spectrum, where \( d \) is the slit separation and \( \lambda \) is the electron or photon wavelength.

Both patterns build up statistically as individual particles fall on the detector. This can be observed for photons or electrons—for now, let us concentrate on electrons. You might imagine that the electrons are interfering with one another as any waves do. To test this, you can lower the intensity until there is never more than one electron between the slits and the screen. The same interference pattern builds up! This implies that a particle’s probability distribution spans both slits, and the particles actually interfere with themselves. Does this also mean that an electron goes through both slits? An electron is a basic unit of matter that is not divisible. But it is a fair question, and so we should look to see if the electron traverses one slit or the other, or both. One possibility is to have coils around the slits that detect charges moving through them. What is observed is that an electron always goes through one slit or the other; it does not split to go through both. But there is a catch. If you determine that the electron went through one of the slits, you no longer get a double-slit pattern—instead, you get single slit interference. There is no escape by using another method of determining which slit the electron went through. Knowing the particle went through one slit forces a single-slit pattern. If you do not observe which slit the electron goes through, you obtain a double-slit pattern.

Heisenberg Uncertainty

Heisenberg Uncertainty

Heisenberg Uncertainty

Heisenberg Uncertainty

How does knowing which slit the electron passed through change the pattern? The answer is fundamentally important—measurement affects the system being observed. Information can be lost, and in some cases it is impossible to measure two physical quantities simultaneously to exact precision. For example, you can measure the position of a moving electron by scattering light or other electrons from it. Those probes have momentum themselves, and by scattering from the electron, they change its momentum in a manner that loses information. There is a limit to absolute knowledge, even in principle.
It was Werner Heisenberg who first stated this limit to knowledge in 1929 as a result of his work on quantum mechanics and the wave characteristics of all particles. (See Figure 29.27). Specifically, consider simultaneously measuring the position and momentum of an electron (it could be any particle). There is an uncertainty in position $\Delta x$ that is approximately equal to the wavelength of the particle. That is,

$$\Delta x \approx \lambda.$$  \hspace{1cm} (29.40)

As discussed above, a wave is not located at one point in space. If the electron’s position is measured repeatedly, a spread in locations will be observed, implying an uncertainty in position $\Delta x$. To detect the position of the particle, we must interact with it, such as having it collide with a detector. In the collision, the particle will lose momentum. This change in momentum could be anywhere from close to zero to the total momentum of the particle, $p = h/\lambda$. It is not possible to tell how much momentum will be transferred to a detector, and so there is an uncertainty in momentum $\Delta p$, too. In fact, the uncertainty in momentum may be as large as the momentum itself, which in equation form means that

$$\Delta p \approx \frac{h}{\lambda}.$$  \hspace{1cm} (29.41)

The uncertainty in position can be reduced by using a shorter-wavelength electron, since $\Delta x \approx \lambda$. But shortening the wavelength increases the uncertainty in momentum, since $\Delta p \approx \frac{h}{\lambda}$. Conversely, the uncertainty in momentum can be reduced by using a longer-wavelength electron, but this increases the uncertainty in position. Mathematically, you can express this trade-off by multiplying the uncertainties. The wavelength cancels, leaving

$$\Delta x \Delta p \approx h.$$  \hspace{1cm} (29.42)

So if one uncertainty is reduced, the other must increase so that their product is $\approx h$.

With the use of advanced mathematics, Heisenberg showed that the best that can be done in a simultaneous measurement of position and momentum is

$$\Delta x \Delta p \geq \frac{h}{4\pi}.$$  \hspace{1cm} (29.43)

This is known as the Heisenberg uncertainty principle. It is impossible to measure position $x$ and momentum $p$ simultaneously with uncertainties $\Delta x$ and $\Delta p$ that multiply to be less than $h/4\pi$. Neither uncertainty can be zero. Neither uncertainty can become small without the other becoming large. A small wavelength allows accurate position measurement, but it increases the momentum of the probe to the point that it further disturbs the momentum of a system being measured. For example, if an electron is scattered from an atom and has a wavelength small enough to detect the position of electrons in the atom, its momentum can knock the electrons from their orbits in a manner that loses information about their original motion. It is therefore impossible to follow an electron in its orbit around an atom. If you measure the electron’s position, you will find it in a definite location, but the atom will be disrupted. Repeated measurements on identical atoms will produce interesting probability distributions for electrons around the atom, but they will not produce motion information. The probability distributions are referred to as electron clouds or orbitals. The shapes of these orbitals are often shown in general chemistry texts and are discussed in The Wave Nature of Matter Causes Quantization.

### Example 29.8 Heisenberg Uncertainty Principle in Position and Momentum for an Atom

(a) If the position of an electron in an atom is measured to an accuracy of 0.0100 nm, what is the electron’s uncertainty in velocity? (b) If the electron has this velocity, what is its kinetic energy in eV?
Strategy
The uncertainty in position is the accuracy of the measurement, or \( \Delta x = 0.0100 \text{ nm} \). Thus the smallest uncertainty in momentum \( \Delta p \) can be calculated using \( \Delta x \Delta p \geq \hbar/4\pi \). Once the uncertainty in momentum \( \Delta p \) is found, the uncertainty in velocity can be found from \( \Delta p = m \Delta v \).

Solution for (a)

Using the equals sign in the uncertainty principle to express the minimum uncertainty, we have

\[ \Delta x \Delta p = \frac{\hbar}{4\pi} \quad (29.44) \]

Solving for \( \Delta p \) and substituting known values gives

\[ \Delta p = \frac{\hbar}{4\pi \Delta x} = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{4\pi(1.00 \times 10^{-11} \text{ m})} = 5.28 \times 10^{-24} \text{ kg} \cdot \text{m/s} \]

Thus,

\[ \Delta p = 5.28 \times 10^{-24} \text{ kg} \cdot \text{m/s} = m \Delta v \quad (29.46) \]

Solving for \( \Delta v \) and substituting the mass of an electron gives

\[ \Delta v = \frac{\Delta p}{m} = \frac{5.28 \times 10^{-24} \text{ kg} \cdot \text{m/s}}{9.11 \times 10^{-31} \text{ kg}} = 5.79 \times 10^6 \text{ m/s} \]

Solution for (b)

Although large, this velocity is not highly relativistic, and so the electron’s kinetic energy is

\[ KE_e = \frac{1}{2}mv^2 \]

\[ = \frac{1}{2}(9.11 \times 10^{-31} \text{ kg})(5.79 \times 10^6 \text{ m/s})^2 \]

\[ = (1.53 \times 10^{-17} \text{ J}) \left( \frac{1 \text{ eV}}{1.60 \times 10^{-19} \text{ J}} \right) = 95.5 \text{ eV} \]

Discussion
Since atoms are roughly 0.1 nm in size, knowing the position of an electron to 0.0100 nm localizes it reasonably well inside the atom. This would be like being able to see details one-tenth the size of the atom. But the consequent uncertainty in velocity is large. You certainly could not follow it very well if its velocity is so uncertain. To get a further idea of how large the uncertainty in velocity is, we assumed the velocity of the electron was equal to its uncertainty and found this gave a kinetic energy of 95.5 eV. This is significantly greater than the typical energy difference between levels in atoms (see Table 29.1), so that it is impossible to get a meaningful energy for the electron if we know its position even moderately well.

Why don’t we notice Heisenberg’s uncertainty principle in everyday life? The answer is that Planck’s constant is very small. Thus the lower limit in the uncertainty of measuring the position and momentum of large objects is negligible. We can detect sunlight reflected from Jupiter and follow the planet in its orbit around the Sun. The reflected sunlight alters the momentum of Jupiter and creates an uncertainty in its momentum, but this is totally negligible compared with Jupiter’s huge momentum. The correspondence principle tells us that the predictions of quantum mechanics become indistinguishable from classical physics for large objects, which is the case here.

Heisenberg Uncertainty for Energy and Time

There is another form of Heisenberg’s uncertainty principle for simultaneous measurements of energy and time. In equation form,

\[ \Delta E \Delta t \geq \frac{\hbar}{4\pi} \quad (29.49) \]

where \( \Delta E \) is the uncertainty in energy and \( \Delta t \) is the uncertainty in time. This means that within a time interval \( \Delta t \), it is not possible to measure energy precisely—there will be an uncertainty \( \Delta E \) in the measurement. In order to measure energy more precisely (to make \( \Delta E \) smaller), we must increase \( \Delta t \). This time interval may be the amount of time we take to make the measurement, or it could be the amount of time a particular state exists, as in the next Example 29.9.

Example 29.9 Heisenberg Uncertainty Principle for Energy and Time for an Atom

An atom in an excited state temporarily stores energy. If the lifetime of this excited state is measured to be \( 1.0 \times 10^{-10} \text{ s} \), what is the minimum uncertainty in the energy of the state in eV?

Strategy
The minimum uncertainty in energy $\Delta E$ is found by using the equals sign in $\Delta E \Delta t \geq \hbar/4\pi$ and corresponds to a reasonable choice for the uncertainty in time. The largest the uncertainty in time can be is the full lifetime of the excited state, or $\Delta t = 1.0 \times 10^{-10}$ s.

**Solution**

Solving the uncertainty principle for $\Delta E$ and substituting known values gives

$$\Delta E = \frac{\hbar}{4\pi \Delta t} = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{4\pi (1.0 \times 10^{-10} \text{ s})} = 5.3 \times 10^{-25} \text{ J}.$$  (29.50)

Now converting to eV yields

$$\Delta E = (5.3 \times 10^{-25} \text{ J}) \left( \frac{1 \text{ eV}}{1.6 \times 10^{-19} \text{ J}} \right) = 3.3 \times 10^{-6} \text{ eV}.$$  (29.51)

**Discussion**

The lifetime of $10^{-10}$ s is typical of excited states in atoms—on human time scales, they quickly emit their stored energy. An uncertainty in energy of only a few millionths of an eV results. This uncertainty is small compared with typical excitation energies in atoms, which are on the order of 1 eV. So here the uncertainty principle limits the accuracy with which we can measure the lifetime and energy of such states, but not very significantly.

The uncertainty principle for energy and time can be of great significance if the lifetime of a system is very short. Then $\Delta t$ is very small, and $\Delta E$ is consequently very large. Some nuclei and exotic particles have extremely short lifetimes (as small as $10^{-25}$ s), causing uncertainties in energy as great as many GeV ($10^{9}$ eV). Stored energy appears as increased rest mass, and so this means that there is significant uncertainty in the rest mass of short-lived particles. When measured repeatedly, a spread of masses or decay energies are obtained. The spread is $\Delta E$. You might ask whether this uncertainty in energy could be avoided by not measuring the lifetime. The answer is no. Nature knows the lifetime, and so its brevity affects the energy of the particle. This is so well established experimentally that the uncertainty in decay energy is used to calculate the lifetime of short-lived states. Some nuclei and particles are so short-lived that it is difficult to measure their lifetime. But if their decay energy can be measured, its spread is $\Delta E$, and this is used in the uncertainty principle ($\Delta E \Delta t \geq \hbar/4\pi$) to calculate the lifetime $\Delta t$.

There is another consequence of the uncertainty principle for energy and time. If energy is uncertain by $\Delta E$, then conservation of energy can be violated by $\Delta E$ for a time $\Delta t$. Neither the physicist nor nature can tell that conservation of energy has been violated, if the violation is temporary and smaller than the uncertainty in energy. While this sounds innocuous enough, we shall see in later chapters that it allows the temporary creation of matter from nothing and has implications for how nature transmits forces over very small distances.

Finally, note that in the discussion of particles and waves, we have stated that individual measurements produce precise or particle-like results. A definite position is determined each time we observe an electron, for example. But repeated measurements produce a spread in values consistent with wave characteristics. The great theoretical physicist Richard Feynman (1918–1988) commented, "What there are, are particles." When you observe enough of them, they distribute themselves as you would expect for a wave phenomenon. However, what there are as they travel we cannot tell because, when we do try to measure, we affect the traveling.

### 29.8 The Particle-Wave Duality Reviewed

**Particle-wave duality**—the fact that all particles have wave properties—is one of the cornerstones of quantum mechanics. We first came across it in the treatment of photons, those particles of EM radiation that exhibit both particle and wave properties, but not at the same time. Later it was noted that particles of matter have wave properties as well. The dual properties of particles and waves are found for all particles, whether massless like photons, or having a mass like electrons. (See Figure 29.28.)

![Figure 29.28](image-url) On a quantum-mechanical scale (i.e., very small), particles with and without mass have wave properties. For example, both electrons and photons have wavelengths but also behave as particles.

There are many submicroscopic particles in nature. Most have mass and are expected to act as particles, or the smallest units of matter. All these masses have wave properties, with wavelengths given by the de Broglie relationship $\lambda = \hbar / p$. So, too, do combinations of these particles, such as nuclei, atoms, and molecules. As a combination of masses becomes large, particularly if it is large enough to be called macroscopic, its wave nature becomes difficult to observe. This is consistent with our common experience with matter.
Some particles in nature are massless. We have only treated the photon so far, but all massless entities travel at the speed of light, have a wavelength, and exhibit particle and wave behaviors. They have momentum given by a rearrangement of the de Broglie relationship, \( p = \frac{h}{\lambda} \). In large combinations of these massless particles (such large combinations are common only for photons or EM waves), there is mostly wave behavior upon detection, and the particle nature becomes difficult to observe. This is also consistent with experience. (See Figure 29.29.)

![Figure 29.29](image)

The particle-wave duality is a universal attribute. It is another connection between matter and energy. Not only has modern physics been able to describe nature for high speeds and small sizes, it has also discovered new connections and symmetries. There is greater unity and symmetry in nature than was known in the classical era—but they were dreamt of. A beautiful poem written by the English poet William Blake some two centuries ago contains the following four lines:

To see the World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And Eternity in an hour

Integrated Concepts

The problem set for this section involves concepts from this chapter and several others. Physics is most interesting when applied to general situations involving more than a narrow set of physical principles. For example, photons have momentum, hence the relevance of Linear Momentum and Collisions. The following topics are involved in some or all of the problems in this section:

- Dynamics: Newton’s Laws of Motion
- Work, Energy, and Energy Resources
- Linear Momentum and Collisions
- Heat and Heat Transfer Methods
- Electric Potential and Electric Field
- Electric Current, Resistance, and Ohm’s Law
- Wave Optics
- Special Relativity

Problem-Solving Strategy

1. Identify which physical principles are involved.
2. Solve the problem using strategies outlined in the text.

Example 29.10 illustrates how these strategies are applied to an integrated-concept problem.

**Example 29.10 Recoil of a Dust Particle after Absorbing a Photon**

The following topics are involved in this integrated concepts worked example:

<table>
<thead>
<tr>
<th>Table 29.2 Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photons (quantum mechanics)</td>
</tr>
<tr>
<td>Linear Momentum</td>
</tr>
</tbody>
</table>

A 550-nm photon (visible light) is absorbed by a 1.00-\(\mu\)g particle of dust in outer space. (a) Find the momentum of such a photon. (b) What is the recoil velocity of the particle of dust, assuming it is initially at rest?

**Strategy Step 1**

To solve an integrated-concept problem, such as those following this example, we must first identify the physical principles involved and identify the chapters in which they are found. Part (a) of this example asks for the momentum of a photon, a topic of the present chapter. Part (b) considers recoil following a collision, a topic of Linear Momentum and Collisions.

**Strategy Step 2**

The following solutions to each part of the example illustrate how specific problem-solving strategies are applied. These involve identifying knowns and unknowns, checking to see if the answer is reasonable, and so on.

**Solution for (a)**

The momentum of a photon is related to its wavelength by the equation:
\[ p = \frac{\hbar}{\lambda} \quad (29.52) \]

Entering the known value for Planck's constant \( \hbar \) and given the wavelength \( \lambda \), we obtain

\[
p = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{550 \times 10^{-9} \text{ m}} = 1.21 \times 10^{-27} \text{ kg} \cdot \text{m/s}. \quad (29.53)
\]

**Discussion for (a)**

This momentum is small, as expected from discussions in the text and the fact that photons of visible light carry small amounts of energy and momentum compared with those carried by macroscopic objects.

**Solution for (b)**

Conservation of momentum in the absorption of this photon by a grain of dust can be analyzed using the equation:

\[
p_1 + p_2 = p'_1 + p'_2 (F_{\text{net}} = 0). \quad (29.54)
\]

The net external force is zero, since the dust is in outer space. Let 1 represent the photon and 2 the dust particle. Before the collision, the dust is at rest (relative to some observer); after the collision, there is no photon (it is absorbed). So conservation of momentum can be written

\[
p_1 = p'_2 = mv, \quad (29.55)
\]

where \( p_1 \) is the photon momentum before the collision and \( p'_2 \) is the dust momentum after the collision. The mass and recoil velocity of the dust are \( m \) and \( v \), respectively. Solving this for \( v \), the requested quantity, yields

\[
v = \frac{p}{m}, \quad (29.56)
\]

where \( p \) is the photon momentum found in part (a). Entering known values (noting that a microgram is \( 10^{-9} \text{ kg} \)) gives

\[
v = \frac{1.21 \times 10^{-27} \text{ kg} \cdot \text{m/s}}{1.00 \times 10^{-9} \text{ kg}} = 1.21 \times 10^{-18} \text{ m/s}. \quad (29.57)
\]

**Discussion**

The recoil velocity of the particle of dust is extremely small. As we have noted, however, there are immense numbers of photons in sunlight and other macroscopic sources. In time, collisions and absorption of many photons could cause a significant recoil of the dust, as observed in comet tails.

---

**Glossary**

- **atomic spectra**: the electromagnetic emission from atoms and molecules
- **binding energy**: also called the *work function*, the amount of energy necessary to eject an electron from a material
- **blackbody radiation**: the electromagnetic radiation from a blackbody
- **blackbody**: an ideal radiator, which can radiate equally well at all wavelengths
- **bremsstrahlung**: German for *braking radiation*, produced when electrons are decelerated
- **Compton effect**: the phenomenon whereby x rays scattered from materials have decreased energy
- **characteristic x rays**: x rays whose energy depends on the material they were produced in
- **correspondence principle**: in the classical limit (large, slow-moving objects), quantum mechanics becomes the same as classical physics
- **de Broglie wavelength**: the wavelength possessed by a particle of matter, calculated by \( \lambda = \frac{\hbar}{p} \)
- **gamma ray**: also γ-ray; highest-energy photon in the EM spectrum
- **Heisenberg’s uncertainty principle**: a fundamental limit to the precision with which pairs of quantities (momentum and position, and energy and time) can be measured
- **infrared radiation**: photons with energies slightly less than red light
- **ionizing radiation**: radiation that ionizes materials that absorb it
- **microwaves**: photons with wavelengths on the order of a micron (\( \mu \text{m} \))
Planck's constant: \( h = 6.626 \times 10^{-34} \text{ J} \cdot \text{s} \)

**particle-wave duality**: the property of behaving like either a particle or a wave; the term for the phenomenon that all particles have wave characteristics

**photoelectric effect**: the phenomenon whereby some materials eject electrons when light is shined on them

**photon energy**: the amount of energy a photon has; \( E = hf \)

**photon momentum**: the amount of momentum a photon has, calculated by \( p = \frac{h}{\lambda} = \frac{E}{c} \)

**photon**: a quantum, or particle, of electromagnetic radiation

**probability distribution**: the overall spatial distribution of probabilities to find a particle at a given location

**quantized**: the fact that certain physical entities exist only with particular discrete values and not every conceivable value

**quantum mechanics**: the branch of physics that deals with small objects and with the quantization of various entities, especially energy

**ultraviolet radiation**: UV; ionizing photons slightly more energetic than violet light

**uncertainty in energy**: lack of precision or lack of knowledge of precise results in measurements of energy

**uncertainty in momentum**: lack of precision or lack of knowledge of precise results in measurements of momentum

**uncertainty in position**: lack of precision or lack of knowledge of precise results in measurements of position

**uncertainty in time**: lack of precision or lack of knowledge of precise results in measurements of time

**visible light**: the range of photon energies the human eye can detect

**x ray**: EM photon between \( \gamma \)-ray and UV in energy

### Section Summary

#### 29.1 Quantization of Energy
- The first indication that energy is sometimes quantized came from blackbody radiation, which is the emission of EM radiation by an object with an emissivity of 1.
- Planck recognized that the energy levels of the emitting atoms and molecules were quantized, with only the allowed values of \( E = (n + \frac{1}{2})hf \), where \( n \) is any non-negative integer (0, 1, 2, 3, ...).
- \( h \) is Planck's constant, whose value is \( h = 6.626 \times 10^{-34} \text{ J} \cdot \text{s} \).
- Thus, the oscillatory absorption and emission energies of atoms and molecules in a blackbody could increase or decrease only in steps of size \( \Delta E = hf \) where \( f \) is the frequency of the oscillatory nature of the absorption and emission of EM radiation.
- Another indication of energy levels being quantized in atoms and molecules comes from the lines in atomic spectra, which are the EM emissions of individual atoms and molecules.

#### 29.2 The Photoelectric Effect
- The photoelectric effect is the process in which EM radiation ejects electrons from a material.
- Einstein proposed photons to be quanta of EM radiation having energy \( E = hf \), where \( f \) is the frequency of the radiation.
- All EM radiation is composed of photons. As Einstein explained, all characteristics of the photoelectric effect are due to the interaction of individual photons with individual electrons.
- The maximum kinetic energy \( KE_e \) of ejected electrons (photoelectrons) is given by \( KE_e = hf - BE \), where \( hf \) is the photon energy and \( BE \) is the binding energy (or work function) of the electron to the particular material.

#### 29.3 Photon Energies and the Electromagnetic Spectrum
- Photon energy is responsible for many characteristics of EM radiation, being particularly noticeable at high frequencies.
- Photons have both wave and particle characteristics.

#### 29.4 Photon Momentum
- Photons have momentum, given by \( p = \frac{h}{\lambda} \), where \( \lambda \) is the photon wavelength.
- Photon energy and momentum are related by \( p = \frac{E}{c} \), where \( E = hf = hc/\lambda \) for a photon.

#### 29.5 The Particle-Wave Duality
- EM radiation can behave like either a particle or a wave.
- This is termed particle-wave duality.
29.6 The Wave Nature of Matter
- Particles of matter also have a wavelength, called the de Broglie wavelength, given by \( \lambda = \frac{h}{p} \), where \( p \) is momentum.
- Matter is found to have the same interference characteristics as any other wave.

29.7 Probability: The Heisenberg Uncertainty Principle
- Matter is found to have the same interference characteristics as any other wave.
- There is now a probability distribution for the location of a particle rather than a definite position.
- Another consequence of the wave character of all particles is the Heisenberg uncertainty principle, which limits the precision with which certain physical quantities can be known simultaneously. For position and momentum, the uncertainty principle is \( \Delta x \Delta p \geq \frac{h}{4\pi} \), where \( \Delta x \) is the uncertainty in position and \( \Delta p \) is the uncertainty in momentum.
- For energy and time, the uncertainty principle is \( \Delta E \Delta t \geq \frac{h}{4\pi} \) where \( \Delta E \) is the uncertainty in energy and \( \Delta t \) is the uncertainty in time.
- These small limits are fundamentally important on the quantum-mechanical scale.

29.8 The Particle-Wave Duality Reviewed
- The particle-wave duality refers to the fact that all particles—those with mass and those without mass—have wave characteristics.
- This is a further connection between mass and energy.

Conceptual Questions

29.1 Quantization of Energy
1. Give an example of a physical entity that is quantized. State specifically what the entity is and what the limits are on its values.
2. Give an example of a physical entity that is not quantized, in that it is continuous and may have a continuous range of values.
3. What aspect of the blackbody spectrum forced Planck to propose quantization of energy levels in its atoms and molecules?
4. If Planck’s constant were large, say \( 10^{34} \) times greater than it is, we would observe macroscopic entities to be quantized. Describe the motions of a child’s swing under such circumstances.
5. Why don’t we notice quantization in everyday events?

29.2 The Photoelectric Effect
6. Is visible light the only type of EM radiation that can cause the photoelectric effect?
7. Which aspects of the photoelectric effect cannot be explained without photons? Which can be explained without photons? Are the latter inconsistent with the existence of photons?
8. Is the photoelectric effect a direct consequence of the wave character of EM radiation or of the particle character of EM radiation? Explain briefly.
9. Insulators (nonmetals) have a higher BE than metals, and it is more difficult for photons to eject electrons from insulators. Discuss how this relates to the free charges in metals that make them good conductors.
10. If you pick up and shake a piece of metal that has electrons in it free to move as a current, no electrons fall out. Yet if you heat the metal, electrons can be boiled off. Explain both of these facts as they relate to the amount and distribution of energy involved with shaking the object as compared with heating it.

29.3 Photon Energies and the Electromagnetic Spectrum
11. Why are UV, x rays, and \( \gamma \) rays called ionizing radiation?
12. How can treating food with ionizing radiation help keep it from spoiling? UV is not very penetrating. What else could be used?
13. Some television tubes are CRTs. They use an approximately 30-kV accelerating potential to send electrons to the screen, where the electrons stimulate phosphors to emit the light that forms the pictures we watch. Would you expect x rays also to be created?
14. Tanning salons use “safe” UV with a longer wavelength than some of the UV in sunlight. This “safe” UV has enough photon energy to trigger the tanning mechanism. Is it likely to be able to cause cell damage and induce cancer with prolonged exposure?
15. Your pupils dilate when visible light intensity is reduced. Does wearing sunglasses that lack UV blockers increase or decrease the UV hazard to your eyes? Explain.
16. One could feel heat transfer in the form of infrared radiation from a large nuclear bomb detonated in the atmosphere 75 km from you. However, none of the profusely emitted x rays or \( \gamma \) rays reaches you. Explain.
18. In an x-ray tube, the maximum photon energy is given by \( hf = qV \). Would it be technically more correct to say \( hf = qV + BE \), where BE is the binding energy of electrons in the target anode? Why isn’t the energy stated the latter way?

29.4 Photon Momentum
19. Which formula may be used for the momentum of all particles, with or without mass?
20. Is there any measurable difference between the momentum of a photon and the momentum of matter?
21. Why don't we feel the momentum of sunlight when we are on the beach?

29.6 The Wave Nature of Matter
22. How does the interference of water waves differ from the interference of electrons? How are they analogous?
23. Describe one type of evidence for the wave nature of matter.
24. Describe one type of evidence for the particle nature of EM radiation.

29.7 Probability: The Heisenberg Uncertainty Principle
25. What is the Heisenberg uncertainty principle? Does it place limits on what can be known?

29.8 The Particle-Wave Duality Reviewed
26. In what ways are matter and energy related that were not known before the development of relativity and quantum mechanics?

In what ways are matter and energy related that were not known before the development of relativity and quantum mechanics?
Problems & Exercises

29.1 Quantization of Energy

1. A LiBr molecule oscillates with a frequency of $1.7 \times 10^{13}$ Hz. (a) What is the difference in energy in eV between allowed oscillator states? (b) What is the approximate value of $n$ for a state having an energy of 1.0 eV?

2. The difference in energy between allowed oscillator states in HBr molecules is 0.330 eV. What is the oscillation frequency of this molecule?

3. A physicist is watching a 15-kg orangutan at a zoo swing lazily in a tire at the end of a rope. He (the physicist) notices that each oscillation takes 3.00 s and hypothesizes that the energy is quantized. (a) What is the difference in energy in joules between allowed oscillator states? (b) What is the value of $n$ for a state where the energy is $5.00 \text{ J}$? (c) Can the quantization be observed?

29.2 The Photoelectric Effect

4. What is the longest-wavelength EM radiation that can eject a photoelectron from silver, given that the binding energy is 4.73 eV? Is this in the visible range?

5. Find the longest-wavelength photon that can eject an electron from potassium, given that the binding energy is 2.42 eV. Is this visible EM radiation?

6. What is the binding energy in eV of electrons in magnesium, if the longest-wavelength photon that can eject electrons is 337 nm?

7. Calculate the binding energy in eV of electrons in aluminum, if the longest-wavelength photon that can eject them is 304 nm.

8. What is the maximum kinetic energy in eV of electrons ejected from sodium metal by 450-nm EM radiation, given that the binding energy is $2.28 \text{ eV}$?

9. UV radiation having a wavelength of 120 nm falls on gold metal, to which electrons are bound by 4.82 eV. What is the maximum kinetic energy of the ejected photoelectrons?

10. Violet light of wavelength 400 nm ejects electrons with a maximum kinetic energy of 0.860 eV from sodium metal. What is the binding energy of electrons to sodium metal?

11. UV radiation having a 300-nm wavelength falls on uranium metal, ejecting 0.500-eV electrons. What is the binding energy of electrons to uranium metal?

12. What is the wavelength of EM radiation that ejects 2.00-eV electrons from calcium metal, given that the binding energy is $2.71 \text{ eV}$? What type of EM radiation is this?

13. Find the wavelength of photons that eject 0.100-eV electrons from potassium, given that the binding energy is 2.24 eV. Are these photons visible?

14. What is the maximum velocity of electrons ejected from a material by 80-nm photons, if they are bound to the material by 4.73 eV?

15. Photoelectrons from a material with a binding energy of 2.71 eV are ejected by 420-nm photons. Once ejected, how long does it take these electrons to travel 2.50 cm to a detection device?

16. A laser with a power output of 2.00 mW at a wavelength of 400 nm is projected onto calcium metal. (a) How many photoelectrons per second are ejected? (b) What power is carried away by the electrons, given that the binding energy is 2.31 eV?

17. (a) Calculate the number of photoelectrons per second ejected from a 1.00-mm² area of sodium metal by 500-nm EM radiation having an intensity of $1.30 \text{ kW/m}^2$ (the intensity of sunlight above the Earth’s atmosphere). (b) Given that the binding energy is 2.28 eV, what power is carried away by the electrons? (c) The electrons carry away less power than brought in by the photons. Where does the other power go? How can it be recovered?

18. Unreasonable Results

---

29.3 Photon Energies and the Electromagnetic Spectrum

20. What is the energy in joules and eV of a photon in a radio wave from an AM station that has a 1530-kHz broadcast frequency?

21. (a) Find the energy in joules and eV of photons in radio waves from an FM station that has a 90.0-MHz broadcast frequency. (b) What does this imply about the number of photons per second that the radio station must broadcast?

22. Calculate the frequency in hertz of a $1.00-\text{MeV} \gamma$-ray photon.

23. (a) What is the wavelength of a 1.00-eV photon? (b) Find its frequency in hertz. (c) Identify the type of EM radiation.

24. Do the unit conversions necessary to show that $hc = 1240 \text{ eV \cdot nm}$, as stated in the text.

25. Confirm the statement in the text that the range of photon energies for visible light is 1.63 to 3.26 eV, given that the range of visible wavelengths is 380 to 760 nm.

26. (a) Calculate the energy in eV of an IR photon of frequency $2.00 \times 10^{13}$ Hz. (b) How many of these photons would need to be absorbed simultaneously by a tightly bound molecule to break it apart? (c) What is the energy in eV of a $\gamma$-ray of frequency $3.00 \times 10^{20}$ Hz? (d) How many tightly bound molecules could a single such $\gamma$ ray break apart?

27. Prove that, to three-digit accuracy, $\frac{h}{c} = 4.14 \times 10^{-15}$ eV \cdot s, as stated in the text.

28. (a) What is the maximum energy in eV of photons produced in a CRT using a 25.0-kV accelerating potential, such as a color TV? (b) What is their frequency?

29. What is the accelerating voltage of an x-ray tube that produces x rays with a shortest wavelength of 0.0103 nm?

30. (a) What is the ratio of power outputs by two microwave ovens having frequencies of 950 and 2560 MHz, if they emit the same number of photons per second? (b) What is the ratio of photons per second if they have the same power output?

31. How many photons per second are emitted by the antenna of a microwave oven, if its power output is 1.00 kW at a frequency of 2560 MHz?

32. Some satellites use nuclear power. (a) If such a satellite emits a 1.00-W flux of $\gamma$ rays having an average energy of 0.500 MeV, how many are emitted per second? (b) These $\gamma$ rays affect other satellites. How far away must another satellite be to only receive one $\gamma$ ray per second per square meter?

33. (a) If the power output of a 650-kHz radio station is 50.0 kW, how many photons per second are produced? (b) If the radio waves are broadcast uniformly in all directions, find the number of photons per second per square meter at a distance of 100 km. Assume no reflection from the ground or absorption by the air.

34. How many x-ray photons per second are created by an x-ray tube that produces a flux of x rays having a power of 1.00 W? Assume the average energy per photon is 75.0 keV.
35. (a) How far away must you be from a 650-kHz radio station with power 50.0 kW for there to be only one photon per second per square meter? Assume no reflections or absorption, as if you were in deep outer space. (b) Discuss the implications for detecting intelligent life in other solar systems by detecting their radio broadcasts.

36. Assuming that 10.0% of a 100-W light bulb's energy output is in the visible range (typical for incandescent bulbs) with an average wavelength of 580 nm, and that the photons spread out uniformly and are not absorbed by the atmosphere, how far away would you be if 500 photons per second enter the 3.00-mm diameter pupil of your eye? (This number easily stimulates the retina.)

37. **Construct Your Own Problem**

Consider a laser pen. Construct a problem in which you calculate the number of photons per second emitted by the pen. Among the things to be considered are the laser pen's wavelength and power output. Your instructor may also wish for you to determine the minimum diffraction spreading in the beam and the number of photons per square centimeter the pen can project at some large distance. In this latter case, you will also need to consider the output size of the laser beam, the distance to the object being illuminated, and any absorption or scattering along the way.

### 29.4 Photon Momentum

38. (a) Find the momentum of a 4.00-cm-wavelength microwave photon. (b) Discuss why you expect the answer to (a) to be very small.

39. (a) What is the momentum of a 0.0100-nm-wavelength photon that could detect details of an atom? (b) What is its energy in MeV?

40. (a) What is the wavelength of a photon that has a momentum of $5.00 \times 10^{-29} \text{ kg} \cdot \text{m/s}$? (b) Find its energy in eV.

41. (a) A γ-ray photon has a momentum of $8.00 \times 10^{-21} \text{ kg} \cdot \text{m/s}$. What is its wavelength? (b) Calculate its energy in MeV.

42. (a) Calculate the momentum of a photon having a wavelength of $2.50 \mu \text{m}$. (b) Find the velocity of an electron having the same momentum. (c) What is the kinetic energy of the electron, and how does it compare with that of the photon?

43. Repeat the previous problem for a 10.0-nm-wavelength photon.

44. (a) Calculate the wavelength of a photon that has the same momentum as a proton moving at 1.00% of the speed of light. (b) What is the energy of the photon in MeV? (c) What is the kinetic energy of the proton in MeV?

45. (a) Find the momentum of a 100-keV x-ray photon. (b) Find the equivalent velocity of a neutron with the same momentum. (c) What is the neutron’s kinetic energy in keV?

46. Take the ratio of relativistic rest energy, $E = \gamma mc^2$, to relativistic momentum, $p = \gamma mu$, and show that in the limit that mass approaches zero, you find $E / p = c$.

### 29.6 The Wave Nature of Matter

49. At what velocity will an electron have a wavelength of 1.00 m?

50. What is the wavelength of an electron moving at 3.00% of the speed of light?

51. At what velocity does a proton have a 6.00-fm wavelength (about the size of a nucleus)? Assume the proton is nonrelativistic. (1 femtometer = $10^{-15}$ m.)

52. What is the velocity of a 0.400-kg billiard ball if its wavelength is 7.50 cm (large enough for it to interfere with other billiard balls)?

53. Find the wavelength of a proton moving at 1.00% of the speed of light.

54. Experiments are performed with ultracold neutrons having velocities as small as 1.00 m/s. (a) What is the wavelength of such a neutron? (b) What is its kinetic energy in eV?

55. (a) Find the velocity of a neutron that has a 6.00-fm wavelength (about the size of a nucleus). Assume the neutron is nonrelativistic. (b) What is the neutron’s kinetic energy in MeV?

56. What is the wavelength of an electron accelerated through a 30.0-kV potential, as in a TV tube?

57. What is the kinetic energy of an electron in a TEM having a 0.0100-nm wavelength?

58. (a) Calculate the velocity of an electron that has a wavelength of 1.00 µm. (b) Through what voltage must the electron be accelerated to have this velocity?

59. The velocity of a proton emerging from a Van de Graaff accelerator is 25.0% of the speed of light. (a) What is the proton’s wavelength? (b) What is its kinetic energy, assuming it is nonrelativistic? (c) What was the equivalent voltage through which it was accelerated?

60. The kinetic energy of an electron accelerated in an x-ray tube is 100 keV. Assuming it is nonrelativistic, what is its wavelength?

### 29.7 Probability: The Heisenberg Uncertainty Principle

62. (a) If the position of an electron in a membrane is measured to an accuracy of 1.00 µm, what is the electron’s minimum uncertainty in velocity? (b) If the electron has this velocity, what is its kinetic energy in eV? (c) What are the implications of this energy, comparing it to typical molecular binding energies?

63. (a) If the position of a chlorine ion in a membrane is measured to an accuracy of 1.00 µm, what is its minimum uncertainty in velocity, given its mass is $5.86 \times 10^{-26}$ kg? (b) If the ion has this velocity, what is its kinetic energy in eV, and how does this compare with typical molecular binding energies?

64. Suppose the velocity of an electron in an atom is known to an accuracy of $2.0 \times 10^{-3}$ m/s (reasonably accurate compared with orbital velocities). What is the electron's minimum uncertainty in position, and how does this compare with the approximate 0.1-nm size of the atom?

65. The velocity of a proton in an accelerator is known to an accuracy of 0.250% of the speed of light. (This could be small compared with its velocity.) What is the smallest possible uncertainty in its position?

66. A relatively long-lived excited state of an atom has a lifetime of 3.00 ms. What is the minimum uncertainty in its energy?

67. (a) The lifetime of a highly unstable nucleus is $10^{-20}$ s. What is the smallest uncertainty in its decay energy? (b) Compare this with the rest energy of an electron.
68. The decay energy of a short-lived particle has an uncertainty of 1.0 MeV due to its short lifetime. What is the smallest lifetime it can have?

69. The decay energy of a short-lived nuclear excited state has an uncertainty of 2.0 eV due to its short lifetime. What is the smallest lifetime it can have?

70. What is the approximate uncertainty in the mass of a muon, as determined from its decay lifetime?

71. Derive the approximate form of Heisenberg's uncertainty principle for energy and time, \( \Delta E \Delta t \approx \hbar \), using the following arguments: Since the position of a particle is uncertain by \( \Delta x \approx \lambda \), where \( \lambda \) is the wavelength of the photon used to examine it, there is an uncertainty in the time the photon takes to traverse \( \Delta x \). Furthermore, the photon has an energy related to its wavelength, and it can transfer some or all of this energy to the object being examined. Thus the uncertainty in the energy of the object is also related to \( \lambda \). Find \( \Delta E \) and \( \Delta t \); then multiply them to give the approximate uncertainty principle.

### 29.8 The Particle-Wave Duality Reviewed

#### 72. Integrated Concepts

The 54.0-eV electron in Example 29.7 has a 0.167-nm wavelength. If such electrons are passed through a double slit and have their first maximum at an angle of \( 25.0^\circ \), what is the slit separation \( d \)?

#### 73. Integrated Concepts

An electron microscope produces electrons with a 2.00-pm wavelength. If these are passed through a 1.00-nm single slit, at what angle will the first diffraction minimum be found?

#### 74. Integrated Concepts

A certain heat lamp emits 200 W of mostly IR radiation averaging 1500 nm in wavelength. (a) What is the average photon energy in joules? (b) How many of these photons are required to increase the temperature of a person's shoulder by \( 2.0^\circ \mathrm{C} \), assuming the affected mass is 4.0 kg with a specific heat of 0.83 kcal/kg \cdot \degree \mathrm{C} \). Also assume no other significant heat transfer. (c) How long does this take?

#### 75. Integrated Concepts

On its high power setting, a microwave oven produces 900 W of 2560 MHz microwaves. (a) How many photons per second is this? (b) How many photons are required to increase the temperature of a 0.500-kg mass of pasta by 45.0 \degree \mathrm{C} \), assuming a specific heat of 0.900 kcal/kg \cdot \degree \mathrm{C} \)? Neglect all other heat transfer. (c) How long must the microwave oven wait for their take to be ready?

#### 76. Integrated Concepts

(a) Calculate the amount of microwave energy in joules needed to raise the temperature of 1.00 kg of soup from 20.0 \degree \mathrm{C} \ to 100 \degree \mathrm{C} \). (b) What is the total momentum of all the microwave photons it takes to do this? (c) Calculate the velocity of a 1.00-kg mass with the same momentum. (d) What is the kinetic energy of this mass?

#### 77. Integrated Concepts

(a) What is \( \gamma \) for an electron emerging from the Stanford Linear Accelerator with a total energy of 50.0 GeV? (b) Find its momentum. (c) What is the electron's wavelength?

#### 78. Integrated Concepts

(a) What is \( \gamma \) for a proton having an energy of 1.00 TeV, produced by the Fermilab accelerator? (b) Find its momentum. (c) What is the proton's wavelength?

#### 79. Integrated Concepts

An electron microscope passes 1.00-pm-wavelength electrons through a circular aperture 2.00 \mu m \ in diameter. What is the angle between two just-resolvable point sources for this microscope?

#### 80. Integrated Concepts

(a) Calculate the velocity of electrons that form the same pattern as 450-nm light when passed through a double slit. (b) Calculate the kinetic energy of each and compare them. (c) Would either be easier to generate than the other? Explain.

#### 81. Integrated Concepts

(a) What is the separation between double slits that produces a second-order minimum at \( 45.0^\circ \) for 650-nm light? (b) What slit separation is needed to produce the same pattern for 1.00-keV protons.

#### 82. Integrated Concepts

A laser with a power output of 2.00 mW at a wavelength of 400 nm is projected onto calcium metal. (a) How many electrons per second are ejected? (b) What power is carried away by the electrons, given that the binding energy is 2.31 eV? (c) Calculate the current of ejected electrons. (d) If the photoelectric material is electrically insulated and acts like a 2.00-pF capacitor, how long will current flow before the capacitor voltage stops it?

#### 83. Integrated Concepts

One problem with x rays is that they are not sensed. Calculate the temperature increase of a researcher exposed in a few seconds to a nearly fatal accidental dose of x rays under the following conditions. The energy of the x-ray photons is 200 keV, and \( 4.00 \times 10^{13} \) of them are absorbed per kilogram of tissue, the specific heat of which is 0.830 kcal/kg \cdot \degree \mathrm{C} \). (Note that medical diagnostic x-ray machines cannot produce an intensity this great.)

#### 84. Integrated Concepts

A 1.00-fm photon has a wavelength short enough to detect some information about nuclei. (a) What is the photon momentum? (b) What is its energy in joules and MeV? (c) What is the (relativistic) velocity of an electron with the same momentum? (d) Calculate the electron's kinetic energy.

#### 85. Integrated Concepts

The momentum of light is exactly reversed when reflected straight back from a mirror, assuming negligible recoil of the mirror. Thus the change in momentum is twice the photon momentum. Suppose light of intensity 1.00 kW/m² reflects from a mirror of area 2.00 m². (a) Calculate the energy reflected in 1.00 s. (b) What is the momentum imparted to the mirror? (c) Using the most general form of Newton’s second law, what is the force on the mirror? (d) Does the assumption of no mirror recoil seem reasonable?

#### 86. Integrated Concepts

Sunlight above the Earth’s atmosphere has an intensity of 1.30 kW/m². If this is reflected straight back from a mirror that has only a small recoil, the light’s momentum is exactly reversed, giving the mirror twice the incident momentum. (a) Calculate the force per square meter of mirror. (b) Very low mass mirrors can be constructed in the near weightlessness of space, and attached to a spaceship to sail it. Once done, the average mass per square meter of the spaceship is 0.100 kg. Find the acceleration of the spaceship if all other forces are balanced. (c) How fast is it moving 24 hours later?
Learning Objectives

30.1. Discovery of the Atom
• Describe the basic structure of the atom, the substructure of all matter.

30.2. Discovery of the Parts of the Atom: Electrons and Nuclei
• Describe how electrons were discovered.
• Explain the Millikan oil drop experiment.
• Describe Rutherford's gold foil experiment.
• Describe Rutherford's planetary model of the atom.

30.3. Bohr's Theory of the Hydrogen Atom
• Describe the mysteries of atomic spectra.
• Explain Bohr’s theory of the hydrogen atom.
• Explain Bohr’s planetary model of the atom.
• Illustrate energy state using the energy-level diagram.
• Describe the triumphs and limits of Bohr’s theory.

30.4. X Rays: Atomic Origins and Applications
• Define x-ray tube and its spectrum.
• Show the x-ray characteristic energy.
• Specify the use of x rays in medical observations.
• Explain the use of x rays in CT scanners in diagnostics.

30.5. Applications of Atomic Excitations and De-Excitations
• Define and discuss fluorescence.
• Define metastable.
• Describe how laser emission is produced.
• Explain population inversion.
• Define and discuss holography.

30.6. The Wave Nature of Matter Causes Quantization
• Explain Bohr’s model of atom.
• Define and describe quantization of angular momentum.
• Calculate the angular momentum for an orbit of atom.
• Define and describe the wave-like properties of matter.

30.7. Patterns in Spectra Reveal More Quantization
• State and discuss the Zeeman effect.
• Define orbital magnetic field.
• Define orbital angular momentum.
• Define space quantization.

30.8. Quantum Numbers and Rules
• Define quantum number.
• Calculate angle of angular momentum vector with an axis.
• Define spin quantum number.

30.9. The Pauli Exclusion Principle
• Define the composition of an atom along with its electrons, neutrons, and protons.
• Explain the Pauli exclusion principle and its application to the atom.
• Specify the shell and subshell symbols and their positions.
• Define the position of electrons in different shells of an atom.
• State the position of each element in the periodic table according to shell filling.
Introduction to Atomic Physics

From childhood on, we learn that atoms are a substructure of all things around us, from the air we breathe to the autumn leaves that blanket a forest trail. Invisible to the eye, the existence and properties of atoms are used to explain many phenomena—a theme found throughout this text. In this chapter, we discuss the discovery of atoms and their own substructures; we then apply quantum mechanics to the description of atoms, and their properties and interactions. Along the way, we will find, much like the scientists who made the original discoveries, that new concepts emerge with applications far beyond the boundaries of atomic physics.

30.1 Discovery of the Atom

How do we know that atoms are really there if we cannot see them with our eyes? A brief account of the progression from the proposal of atoms by the Greeks to the first direct evidence of their existence follows.

People have long speculated about the structure of matter and the existence of atoms. The earliest significant ideas to survive are due to the ancient Greeks in the fifth century BCE, especially those of the philosophers Leucippus and Democritus. (There is some evidence that philosophers in both India and China made similar speculations, at about the same time.) They considered the question of whether a substance can be divided without limit into ever smaller pieces. There are only a few possible answers to this question. One is that infinitesimally small subdivision is possible. Another is what Democritus in particular believed—that there is a smallest unit that cannot be further subdivided. Democritus called this the atom. We now know that atoms themselves can be subdivided, but their identity is destroyed in the process, so the Greeks were correct in a respect. The Greeks also felt that atoms were in constant motion, another correct notion.

The Greeks and others speculated about the properties of atoms, proposing that only a few types existed and that all matter was formed as various combinations of these types. The famous proposal that the basic elements were earth, air, fire, and water was brilliant, but incorrect. The Greeks had identified the most common examples of the four states of matter (solid, gas, plasma, and liquid), rather than the basic elements. More than 2000 years passed before observations could be made with equipment capable of revealing the true nature of atoms.

Over the centuries, discoveries were made regarding the properties of substances and their chemical reactions. Certain systematic features were recognized, but similarities between common and rare elements resulted in efforts to transmute them (lead into gold, in particular) for financial gain. Secrecy was endemic. Alchemists discovered and rediscovered many facts but did not make them broadly available. As the Middle Ages ended, alchemy gradually faded, and the science of chemistry arose. It was no longer possible, nor considered desirable, to keep discoveries secret. Collective knowledge grew, and by the beginning of the 19th century, an important fact was well established—the masses of reactants in specific chemical reactions always have a particular mass ratio. This is very strong indirect evidence that there are basic units (atoms and molecules) that have these same mass ratios. The English chemist John Dalton (1766–1844) did much of this work, with significant contributions by the Italian chemist Amedeo Avogadro (1776–1856). It was Avogadro who developed the idea of a fixed number of atoms and molecules in a mole, and this special number is called Avogadro’s number in his honor. The Austrian physicist Johann Josef Loschmidt was the first to measure the value of the constant in 1865 using the kinetic theory of gases.

Patterns and Systematics

The recognition and appreciation of patterns has enabled us to make many discoveries. The periodic table of elements was proposed as an organized summary of the known elements long before all elements had been discovered, and it led to many other discoveries. We shall see in later chapters that patterns in the properties of subatomic particles led to the proposal of quarks as their underlying structure, an idea that is still bearing fruit.

Knowledge of the properties of elements and compounds grew, culminating in the mid-19th-century development of the periodic table of the elements by Dmitri Mendeleev (1834–1907), the great Russian chemist. Mendeleev proposed an ingenious array that highlighted the periodic nature of the properties of elements. Believing in the systematics of the periodic table, he also predicted the existence of then-unknown elements to complete it. Once these elements were discovered and determined to have properties predicted by Mendeleev, his periodic table became universally accepted.

Also during the 19th century, the kinetic theory of gases was developed. Kinetic theory is based on the existence of atoms and molecules in random thermal motion and provides a microscopic explanation of the gas laws, heat transfer, and thermodynamics (see Introduction to Temperature, Kinetic Theory, and the Gas Laws and Introduction to Laws of Thermodynamics). Kinetic theory works so well that it is another strong indication of the existence of atoms. But it is still indirect evidence—individual atoms and molecules had not been observed. There were heated debates about the validity of kinetic theory until direct evidence of atoms was obtained.

The first truly direct evidence of atoms is credited to Robert Brown, a Scottish botanist. In 1827, he noticed that tiny pollen grains suspended in still water moved about in complex paths. This can be observed with a microscope for any small particles in a fluid. The motion is caused by the random thermal motions of fluid molecules colliding with particles in the fluid, and it is now called Brownian motion. (See Figure 30.2) Statistical fluctuations in the numbers of molecules striking the sides of a visible particle cause it to move first this way, then that. Although the molecules cannot be directly observed, their effects on the particle can be. By examining Brownian motion, the size of molecules can be calculated. The smaller and more numerous they are, the smaller the fluctuations in the numbers striking different sides.
Figure 30.2 The position of a pollen grain in water, measured every few seconds under a microscope, exhibits Brownian motion. Brownian motion is due to fluctuations in the number of atoms and molecules colliding with a small mass, causing it to move about in complex paths. This is nearly direct evidence for the existence of atoms, providing a satisfactory alternative explanation cannot be found.

It was Albert Einstein who, starting in his epochal year of 1905, published several papers that explained precisely how Brownian motion could be used to measure the size of atoms and molecules. (In 1905 Einstein created special relativity, proposed photons as quanta of EM radiation, and produced a theory of Brownian motion that allowed the size of atoms to be determined. All of this was done in his spare time, since he worked days as a patent examiner. Any one of these very basic works could have been the crowning achievement of an entire career—yet Einstein did even more in later years.) Their sizes were only approximately known to be \(10^{-10}\) m, based on a comparison of latent heat of vaporization and surface tension made in about 1805 by Thomas Young of double-slit fame and the famous astronomer and mathematician Simon Laplace.

Using Einstein’s ideas, the French physicist Jean-Baptiste Perrin (1870–1942) carefully observed Brownian motion; not only did he confirm Einstein’s theory, he also produced accurate sizes for atoms and molecules. Since molecular weights and densities of materials were well established, knowing atomic and molecular sizes allowed a precise value for Avogadro’s number to be obtained. (If we know how big an atom is, we know how many fit into a certain volume.) Perrin also used these ideas to explain atomic and molecular agitation effects in sedimentation, and he received the 1926 Nobel Prize for his achievements. Most scientists were already convinced of the existence of atoms, but the accurate observation and analysis of Brownian motion was conclusive—it was the first truly direct evidence.

A huge array of direct and indirect evidence for the existence of atoms now exists. For example, it has become possible to accelerate ions (much as electrons are accelerated in cathode-ray tubes) and to detect them individually as well as measure their masses (see More Applications of Magnetism for a discussion of mass spectrometers). Other devices that observe individual atoms, such as the scanning tunneling electron microscope, will be discussed elsewhere. (See Figure 30.3.) All of our understanding of the properties of matter is based on and consistent with the atom. The atom’s substructures, such as electron shells and the nucleus, are both interesting and important. The nucleus in turn has a substructure, as do the particles of which it is composed. These topics, and the question of whether there is a smallest basic structure to matter, will be explored in later parts of the text.

Figure 30.3 Individual atoms can be detected with devices such as the scanning tunneling electron microscope that produced this image of individual gold atoms on a graphite substrate. (credit: Erwin Rossen, Eindhoven University of Technology, via Wikimedia Commons)

30.2 Discovery of the Parts of the Atom: Electrons and Nuclei

Just as atoms are a substructure of matter, electrons and nuclei are substructures of the atom. The experiments that were used to discover electrons and nuclei reveal some of the basic properties of atoms and can be readily understood using ideas such as electrostatic and magnetic force, already covered in previous chapters.

Charges and Electromagnetic Forces

In previous discussions, we have noted that positive charge is associated with nuclei and negative charge with electrons. We have also covered many aspects of the electric and magnetic forces that affect charges. We will now explore the discovery of the electron and nucleus as substructures of the atom and examine their contributions to the properties of atoms.
The Electron

Gas discharge tubes, such as that shown in Figure 30.4, consist of an evacuated glass tube containing two metal electrodes and a rarefied gas. When a high voltage is applied to the electrodes, the gas glows. These tubes were the precursors to today’s neon lights. They were first studied seriously by Heinrich Geissler, a German inventor and glassblower, starting in the 1860s. The English scientist William Crookes, among others, continued to study what for some time were called Crookes tubes, wherein electrons are freed from atoms and molecules in the rarefied gas inside the tube and are accelerated from the cathode (negative) to the anode (positive) by the high potential. These “cathode rays” collide with the gas atoms and molecules and excite them, resulting in the emission of electromagnetic (EM) radiation that makes the electrons’ path visible as a ray that spreads and fades as it moves away from the cathode.

Gas discharge tubes today are most commonly called cathode-ray tubes, because the rays originate at the cathode. Crookes showed that the electrons carry momentum (they can make a small paddle wheel rotate). He also found that their normally straight path is bent by a magnet in the direction expected for a negative charge moving away from the cathode. These were the first direct indications of electrons and their charge.

The English physicist J. J. Thomson (1856–1940) improved and expanded the scope of experiments with gas discharge tubes. (See Figure 30.5 and Figure 30.6.) He verified the negative charge of the cathode rays with both magnetic and electric fields. Additionally, he collected the rays in a metal cup and found an excess of negative charge. Thomson was also able to measure the ratio of the charge of the electron to its mass, \( \frac{q_e}{m_e} \)—an important step to finding the actual values of both \( q_e \) and \( m_e \). Figure 30.7 shows a cathode-ray tube, which produces a narrow beam of electrons that passes through charging plates connected to a high-voltage power supply. An electric field \( \mathbf{E} \) is produced between the charging plates, and the cathode-ray tube is placed between the poles of a magnet so that the electric field \( \mathbf{E} \) is perpendicular to the magnetic field \( \mathbf{B} \) of the magnet. These fields, being perpendicular to each other, produce opposing forces on the electrons. As discussed for mass spectrometers in More Applications of Magnetism, if the net force due to the fields vanishes, then the velocity of the charged particle is \( v = \frac{E}{B} \). In this manner, Thomson determined the velocity of the electrons and then moved the beam up and down by adjusting the electric field.

Figure 30.4 A gas discharge tube glows when a high voltage is applied to it. Electrons emitted from the cathode are accelerated toward the anode; they excite atoms and molecules in the gas, which glow in response. Once called Geissler tubes and later Crookes tubes, they are now known as cathode-ray tubes (CRTs) and are found in older TVs, computer screens, and x-ray machines. When a magnetic field is applied, the beam bends in the direction expected for negative charge. (credit: Paul Downey, Flickr)

The English physicist J. J. Thomson (1856–1940) improved and expanded the scope of experiments with gas discharge tubes. (See Figure 30.5 and Figure 30.6.) He verified the negative charge of the cathode rays with both magnetic and electric fields. Additionally, he collected the rays in a metal cup and found an excess of negative charge. Thomson was also able to measure the ratio of the charge of the electron to its mass, \( \frac{q_e}{m_e} \)—an important step to finding the actual values of both \( q_e \) and \( m_e \). Figure 30.7 shows a cathode-ray tube, which produces a narrow beam of electrons that passes through charging plates connected to a high-voltage power supply. An electric field \( \mathbf{E} \) is produced between the charging plates, and the cathode-ray tube is placed between the poles of a magnet so that the electric field \( \mathbf{E} \) is perpendicular to the magnetic field \( \mathbf{B} \) of the magnet. These fields, being perpendicular to each other, produce opposing forces on the electrons. As discussed for mass spectrometers in More Applications of Magnetism, if the net force due to the fields vanishes, then the velocity of the charged particle is \( v = \frac{E}{B} \). In this manner, Thomson determined the velocity of the electrons and then moved the beam up and down by adjusting the electric field.

Figure 30.5 J. J. Thomson (credit: www.firstworldwar.com, via Wikimedia Commons)
To see how the amount of deflection is used to calculate $\frac{q_e}{m_e}$, note that the deflection is proportional to the electric force on the electron:

$$F = q_eE.$$  \hspace{1cm} (30.1)

But the vertical deflection is also related to the electron’s mass, since the electron’s acceleration is

$$a = \frac{F}{m_e}.$$  \hspace{1cm} (30.2)

The value of $F$ is not known, since $q_e$ was not yet known. Substituting the expression for electric force into the expression for acceleration yields

$$a = \frac{F}{m_e} = \frac{q_eE}{m_e}.$$  \hspace{1cm} (30.3)

Gathering terms, we have

$$\frac{q_e}{m_e} = \frac{a}{E}.$$  \hspace{1cm} (30.4)

The deflection is analyzed to get $a$, and $E$ is determined from the applied voltage and distance between the plates; thus, $\frac{q_e}{m_e}$ can be determined.

With the velocity known, another measurement of $\frac{q_e}{m_e}$ can be obtained by bending the beam of electrons with the magnetic field. Since $F_{\text{mag}} = q_evB = m_ea$, we have $\frac{q_e}{m_e} = a/vB$. Consistent results are obtained using magnetic deflection.

What is so important about $\frac{q_e}{m_e}$, the ratio of the electron’s charge to its mass? The value obtained is

$$\frac{q_e}{m_e} = -1.76 \times 10^{11} \text{ C/kg (electron).}$$  \hspace{1cm} (30.5)

This is a huge number, as Thomson realized, and it implies that the electron has a very small mass. It was known from electroplating that about $10^8$ C/kg is needed to plate a material, a factor of about 1000 less than the charge per kilogram of electrons. Thomson went on to do the same experiment for positively charged hydrogen ions (now known to be bare protons) and found a charge per kilogram about 1000 times smaller than that for the electron, implying that the proton is about 1000 times more massive than the electron. Today, we know more precisely that

$$\frac{q_p}{m_p} = 9.58 \times 10^7 \text{ C/kg (proton),}$$  \hspace{1cm} (30.6)

where $q_p$ is the charge of the proton and $m_p$ is its mass. This ratio (to four significant figures) is 1836 times less charge per kilogram than for the electron. Since the charges of electrons and protons are equal in magnitude, this implies $m_p = 1836m_e$.

Thomson performed a variety of experiments using differing gases in discharge tubes and employing other methods, such as the photoelectric effect, for freeing electrons from atoms. He always found the same properties for the electron, proving it to be an independent particle. For his work, the
important pieces of which he began to publish in 1897, Thomson was awarded the 1906 Nobel Prize in Physics. In retrospect, it is difficult to appreciate how astonishing it was to find that the atom has a substructure. Thomson himself said, “It was only when I was convinced that the experiment left no escape from it that I published my belief in the existence of bodies smaller than atoms.”

Thomson attempted to measure the charge of individual electrons, but his method could determine its charge only to the order of magnitude expected.

Since Faraday’s experiments with electroplating in the 1830s, it had been known that about 100,000 C per mole was needed to plate singly ionized ions. Dividing this by the number of ions per mole (that is, by Avogadro’s number), which was approximately known, the charge per ion was calculated to be about \(1.6 \times 10^{-19}\) C, close to the actual value.

An American physicist, Robert Millikan (1868–1953) (see Figure 30.8), decided to improve upon Thomson’s experiment for measuring \(q_e\) and was eventually forced to try another approach, which is now a classic experiment performed by students. The Millikan oil drop experiment is shown in Figure 30.9.

Figure 30.8 Robert Millikan (credit: Unknown Author, via Wikimedia Commons)

Figure 30.9 The Millikan oil drop experiment produced the first accurate direct measurement of the charge on electrons, one of the most fundamental constants in nature. Fine drops of oil become charged when sprayed. Their movement is observed between metal plates with a potential applied to oppose the gravitational force. The balance of gravitational and electric forces allows the calculation of the charge on a drop. The charge is found to be quantized in units of \(-1.6 \times 10^{-19}\) C, thus determining directly the charge of the excess and missing electrons on the oil drops.

In the Millikan oil drop experiment, fine drops of oil are sprayed from an atomizer. Some of these are charged by the process and can then be suspended between metal plates by a voltage between the plates. In this situation, the weight of the drop is balanced by the electric force:

\[
m_{\text{drop}} g = q_e E
\]

(30.7)

The electric field is produced by the applied voltage, hence, \(E = V / d\), and \(V\) is adjusted to just balance the drop’s weight. The drops can be seen as points of reflected light using a microscope, but they are too small to directly measure their size and mass. The mass of the drop is determined by observing how fast it falls when the voltage is turned off. Since air resistance is very significant for these submicroscopic drops, the more massive drops fall faster than the less massive, and sophisticated sedimentation calculations can reveal their mass. Oil is used rather than water, because it does not readily evaporate, and so mass is nearly constant. Once the mass of the drop is known, the charge of the electron is given by rearranging the previous equation:

\[
q = q_{e} = \frac{m_{\text{drop}} g d}{E}
\]

(30.8)

where \(d\) is the separation of the plates and \(V\) is the voltage that holds the drop motionless. (The same drop can be observed for several hours to see that it really is motionless.) By 1913 Millikan had measured the charge of the electron \(q_e\) to an accuracy of 1%, and he improved this by a factor...
of 10 within a few years to a value of \(-1.60\times10^{-19}\) C. He also observed that all charges were multiples of the basic electron charge and that sudden changes could occur in which electrons were added or removed from the drops. For this very fundamental direct measurement of \(q_e\) and for his studies of the photoelectric effect, Millikan was awarded the 1923 Nobel Prize in Physics.

With the charge of the electron known and the charge-to-mass ratio known, the electron’s mass can be calculated. It is

\[
m = \frac{q_e}{(4\pi\varepsilon_0)^{1/2}}
\]

Substituting known values yields

\[
m_e = \frac{-1.60\times10^{-19}}{-1.76\times10^{11}} \text{ C/kg}
\]

or

\[
m_e = 9.11\times10^{-31} \text{ kg (electron’s mass)},
\]

where the round-off errors have been corrected. The mass of the electron has been verified in many subsequent experiments and is now known to an accuracy of better than one part in one million. It is an incredibly small mass and remains the smallest known mass of any particle that has mass. (Some particles, such as photons, are massless and cannot be brought to rest, but travel at the speed of light.) A similar calculation gives the masses of other particles, including the proton. To three digits, the mass of the proton is now known to be

\[
m_p = 1.67\times10^{-27} \text{ kg (proton’s mass)},
\]

which is nearly identical to the mass of a hydrogen atom. What Thomson and Millikan had done was to prove the existence of one substructure of atoms, the electron, and further to show that it had only a tiny fraction of the mass of an atom. The nucleus of an atom contains most of its mass, and the nature of the nucleus was completely unanticipated.

Another important characteristic of quantum mechanics was also beginning to emerge. All electrons are identical to one another. The charge and mass of electrons are not average values; rather, they are unique values that all electrons have. This is true of other fundamental entities at the submicroscopic level. All protons are identical to one another, and so on.

The Nucleus

Here, we examine the first direct evidence of the size and mass of the nucleus. In later chapters, we will examine many other aspects of nuclear physics, but the basic information on nuclear size and mass is so important to understanding the atom that we consider it here.

Nuclear radioactivity was discovered in 1896, and it was soon the subject of intense study by a number of the best scientists in the world. Among them was New Zealander Lord Ernest Rutherford, who made numerous fundamental discoveries and earned the title of “father of nuclear physics.” Born in Nelson, Rutherford did his postgraduate studies at the Cavendish Laboratories in England before taking up a position at McGill University in Canada where he did the work that earned him a Nobel Prize in Chemistry in 1908. In the area of atomic and nuclear physics, there is much overlap between chemistry and physics, with physics providing the fundamental enabling theories. He returned to England in later years and had six future Nobel Prize winners as students. Rutherford used nuclear radiation to directly examine the size and mass of the atomic nucleus. The experiment he devised is shown in Figure 30.10. A radioactive source that emits alpha radiation was placed in a lead container with a hole in one side to produce a beam of alpha particles, which are a type of ionizing radiation ejected by the nuclei of a radioactive source. A thin gold foil was placed in the beam, and the scattering of the alpha particles was observed by the glow they caused when they struck a phosphor screen.

![Figure 30.10](https://example.com/figure30.10.png)

Figure 30.10 Rutherford’s experiment gave direct evidence for the size and mass of the nucleus by scattering alpha particles from a thin gold foil. Alpha particles with energies of about 5 MeV are emitted from a radioactive source (which is a small metal container in which a specific amount of a radioactive material is sealed), are collimated into a beam, and fall upon the foil. The number of particles that penetrate the foil or scatter to various angles indicates that gold nuclei are very small and contain nearly all of the gold atom’s mass. This is particularly indicated by the alpha particles that scatter to very large angles, much like a soccer ball bouncing off a goalie’s head.

Alpha particles were known to be the doubly charged positive nuclei of helium atoms that had kinetic energies on the order of 5 MeV when emitted in nuclear decay, which is the disintegration of the nucleus of an unstable nuclide by the spontaneous emission of charged particles. These particles interact with matter mostly via the Coulomb force, and the manner in which they scatter from nuclei can reveal nuclear size and mass. This is analogous to observing how a bowling ball is scattered by an object you cannot see directly. Because the alpha particle’s energy is so large compared with the typical energies associated with atoms (MeV versus eV), you would expect the alpha particles to simply crash through a thin...
foil much like a supersonic bowling ball would crash through a few dozen rows of bowling pins. Thomson had envisioned the atom to be a small sphere in which equal amounts of positive and negative charge were distributed evenly. The incident massive alpha particles would suffer only small deflections in such a model. Instead, Rutherford and his collaborators found that alpha particles occasionally were scattered to large angles, some even back in the direction from which they came! Detailed analysis using conservation of momentum and energy—particularly of the small number that came straight back—implied that gold nuclei are very small compared with the size of a gold atom, contain almost all of the atom’s mass, and are tightly bound. Since the gold nucleus is several times more massive than the alpha particle, a head-on collision would scatter the alpha particle straight back toward the source. In addition, the smaller the nucleus, the fewer alpha particles that would hit one head on.

Although the results of the experiment were published by his colleagues in 1909, it took Rutherford two years to convince himself of their meaning. Like Thomson before him, Rutherford was reluctant to accept such radical results. Nature on a small scale is so unlike our classical world that even those at the forefront of discovery are sometimes surprised. Rutherford later wrote: “It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you. On consideration, I realized that this scattering backwards ... [meant] ... the greatest part of the mass of the atom was concentrated in a tiny nucleus.” In 1911, Rutherford published his analysis together with a proposed model of the atom. The size of the nucleus was determined to be about $10^{-15} \text{ m}$, or 100,000 times smaller than the atom. This implies a huge density, on the order of $10^{15} \text{ g/cm}^3$, vastly unlike any macroscopic matter. Also implied is the existence of previously unknown nuclear forces to counteract the huge repulsive Coulomb forces among the positive charges in the nucleus. Huge forces would also be consistent with the large energies emitted in nuclear radiation.

The small size of the nucleus also implies that the atom is mostly empty inside. In fact, in Rutherford’s experiment, most alphas went straight through the gold foil with very little scattering, since electrons have such small masses and since the atom was mostly empty with nothing for the alpha to hit. There were already hints of this at the time Rutherford performed his experiments, since energetic electrons had been observed to penetrate thin foils more easily than expected. Figure 30.11 shows a schematic of the atoms in a thin foil with circles representing the size of the atoms (about $10^{-10} \text{ m}$) and dots representing the nuclei. (The dots are not to scale—if they were, you would need a microscope to see them.) Most alpha particles miss the small nuclei and are only slightly scattered by electrons. Occasionally, (about once in 8000 times in Rutherford’s experiment), an alpha hits a nucleus head-on and is scattered straight backward.

**Figure 30.11** An expanded view of the atoms in the gold foil in Rutherford’s experiment. Circles represent the atoms (about $10^{-10} \text{ m}$ in diameter), while the dots represent the nuclei (about $10^{-15} \text{ m}$ in diameter). To be visible, the dots are much larger than scale. Most alpha particles crash through but are relatively unaffected because of their high energy and the electron’s small mass. Some, however, head straight toward a nucleus and are scattered straight back. A detailed analysis gives the size and mass of the nucleus.

Based on the size and mass of the nucleus revealed by his experiment, as well as the mass of electrons, Rutherford proposed the planetary model of the atom. The planetary model of the atom pictures low-mass electrons orbiting a large-mass nucleus. The sizes of the electron orbits are large compared with the size of the nucleus, with mostly vacuum inside the atom. This picture is analogous to how low-mass planets in our solar system orbit the large-mass Sun at distances large compared with the size of the sun. In the atom, the attractive Coulomb force is analogous to gravitation in the planetary system. (See Figure 30.12.) Note that a model or mental picture is needed to explain experimental results, since the atom is too small to be directly observed with visible light.

**Figure 30.12** Rutherford’s planetary model of the atom incorporates the characteristics of the nucleus, electrons, and the size of the atom. This model was the first to recognize the structure of atoms, in which low-mass electrons orbit a very small, massive nucleus in orbits much larger than the nucleus. The atom is mostly empty and is analogous to our planetary system.

Rutherford’s planetary model of the atom was crucial to understanding the characteristics of atoms, and their interactions and energies, as we shall see in the next few sections. Also, it was an indication of how different nature is from the familiar classical world on the small, quantum mechanical scale. The discovery of a substructure to all matter in the form of atoms and molecules was now being taken a step further to reveal a substructure of atoms that was simpler than the 92 elements then known. We have continued to search for deeper substructures, such as those inside the nucleus, with some success. In later chapters, we will follow this quest in the discussion of quarks and other elementary particles, and we will look at the direction the search seems now to be heading.
How did Rutherford figure out the structure of the atom without being able to see it? Simulate the famous experiment in which he disproved the Plum Pudding model of the atom by observing alpha particles bouncing off atoms and determining that they must have a small core.

**30.3 Bohr’s Theory of the Hydrogen Atom**

The great Danish physicist Niels Bohr (1885–1962) made immediate use of Rutherford’s planetary model of the atom. (Figure 30.14). Bohr became convinced of its validity and spent part of 1912 at Rutherford’s laboratory. In 1913, after returning to Copenhagen, he began publishing his theory of the simplest atom, hydrogen, based on the planetary model of the atom. For decades, many questions had been asked about atomic characteristics. From their sizes to their spectra, much was known about atoms, but little had been explained in terms of the laws of physics. Bohr’s theory explained the atomic spectrum of hydrogen and established new and broadly applicable principles in quantum mechanics.

Mysteries of Atomic Spectra

As noted in Quantization of Energy, the energies of some small systems are quantized. Atomic and molecular emission and absorption spectra have been known for over a century to be discrete (or quantized). (See Figure 30.15.) Maxwell and others had realized that there must be a connection between the spectrum of an atom and its structure, something like the resonant frequencies of musical instruments. But, in spite of years of efforts by many great minds, no one had a workable theory. (It was a running joke that any theory of atomic and molecular spectra could be destroyed by throwing a book of data at it, so complex were the spectra.) Following Einstein’s proposal of photons with quantized energies directly proportional to their wavelengths, it became even more evident that electrons in atoms can exist only in discrete orbits.
In some cases, it had been possible to devise formulas that described the emission spectra. As you might expect, the simplest atom—hydrogen, with its single electron—has a relatively simple spectrum. The hydrogen spectrum had been observed in the infrared (IR), visible, and ultraviolet (UV), and several series of spectral lines had been observed. (See Figure 30.16.) These series are named after early researchers who studied them in particular depth.

The observed hydrogen-spectrum wavelengths can be calculated using the following formula:

$$\frac{1}{\lambda} = R \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right),$$

where $\lambda$ is the wavelength of the emitted EM radiation and $R$ is the Rydberg constant, determined by the experiment to be

$$R = 1.097 \times 10^7 \text{ m}^{-1}.$$  

The constant $n_f$ is a positive integer associated with a specific series. For the Lyman series, $n_f = 1$; for the Balmer series, $n_f = 2$; for the Paschen series, $n_f = 3$; and so on. The Lyman series is entirely in the UV, while part of the Balmer series is visible with the remainder UV. The Paschen series and all the rest are entirely IR. There are apparently an unlimited number of series, although they lie progressively farther into the infrared and become difficult to observe as $n_f$ increases. The constant $n_i$ is a positive integer, but it must be greater than $n_f$. Thus, for the Balmer series, $n_f = 2$ and $n_i = 3, 4, 5, 6, \ldots$. Note that $n_i$ can approach infinity. While the formula in the wavelengths equation was just a recipe designed to fit data and was not based on physical principles, it did imply a deeper meaning. Balmer first devised the formula for his series alone, and it was later found to describe all the other series by using different values of $n_f$. Bohr was the first to comprehend the deeper meaning. Again, we see the interplay between experiment and theory in physics. Experimentally, the spectra were well established, an equation was found to fit the experimental data, but the theoretical foundation was missing.

![Figure 30.15](http://cnx.org/content/col11406/1.7)

**Example 30.1 Calculating Wave Interference of a Hydrogen Line**

What is the distance between the slits of a grating that produces a first-order maximum for the second Balmer line at an angle of $15^\circ$?

**Strategy and Concept**
For an Integrated Concept problem, we must first identify the physical principles involved. In this example, we need to know (a) the wavelength of light as well as (b) conditions for an interference maximum for the pattern from a double slit. Part (a) deals with a topic of the present chapter, while part (b) considers the wave interference material of Wave Optics.

Solution for (a)

Hydrogen spectrum wavelength. The Balmer series requires that $n_f = 2$. The first line in the series is taken to be for $n_i = 3$, and so the second would have $n_i = 4$.

The calculation is a straightforward application of the wavelength equation. Entering the determined values for $n_f$ and $n_i$ yields

$$\frac{1}{\lambda} = R \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right)$$

$$= \left( 1.097 \times 10^7 \text{ m}^{-1} \right) \left( \frac{1}{2^2} - \frac{1}{4^2} \right)$$

$$= 2.057 \times 10^6 \text{ m}^{-1}.$$

Inverting to find $\lambda$ gives

$$\lambda = \frac{1}{2.057 \times 10^6 \text{ m}^{-1}} = 486 \times 10^{-9} \text{ m}$$

$$= 486 \text{ nm}.$$

Discussion for (a)

This is indeed the experimentally observed wavelength, corresponding to the second (blue-green) line in the Balmer series. More impressive is the fact that the same simple recipe predicts all of the hydrogen spectrum lines, including new ones observed in subsequent experiments. What is nature telling us?

Solution for (b)

Double-slit interference (Wave Optics). To obtain constructive interference for a double slit, the path length difference from two slits must be an integral multiple of the wavelength. This condition was expressed by the equation

$$d \sin \theta = m\lambda,$$  \hspace{1cm} (30.17)

where $d$ is the distance between slits and $\theta$ is the angle from the original direction of the beam. The number $m$ is the order of the interference; $m = 1$ in this example. Solving for $d$ and entering known values yields

$$d = \frac{(1)(486 \text{ nm})}{\sin 15^\circ} = 1.88 \times 10^{-6} \text{ m}.$$  \hspace{1cm} (30.18)

Discussion for (b)

This number is similar to those used in the interference examples of Introduction to Quantum Physics (and is close to the spacing between slits in commonly used diffraction glasses).

Bohr’s Solution for Hydrogen

Bohr was able to derive the formula for the hydrogen spectrum using basic physics, the planetary model of the atom, and some very important new proposals. His first proposal is that only certain orbits are allowed: we say that the orbits of electrons in atoms are quantized. Each orbit has a different energy, and electrons can move to a higher orbit by absorbing energy and drop to a lower orbit by emitting energy. If the orbits are quantized, the amount of energy absorbed or emitted is also quantized, producing discrete spectra. Photon absorption and emission are among the primary methods of transferring energy into and out of atoms. The energies of the photons are quantized, and their energy is explained as being equal to the change in energy of the electron when it moves from one orbit to another. In equation form, this is

$$\Delta E = hf = E_i - E_f,$$  \hspace{1cm} (30.19)

Here, $\Delta E$ is the change in energy between the initial and final orbits, and $hf$ is the energy of the absorbed or emitted photon. It is quite logical (that is, expected from our everyday experience) that energy is involved in changing orbits. A blast of energy is required for the space shuttle, for example, to climb to a higher orbit. What is not expected is that atomic orbits should be quantized. This is not observed for satellites or planets, which can have any orbit given the proper energy. (See Figure 30.17.)
Figure 30.17 The planetary model of the atom, as modified by Bohr, has the orbits of the electrons quantized. Only certain orbits are allowed, explaining why atomic spectra are discrete (quantized). The energy carried away from an atom by a photon comes from the electron dropping from one allowed orbit to another and is thus quantized. This is likewise true for atomic absorption of photons.

Figure 30.18 shows an energy-level diagram, a convenient way to display energy states. In the present discussion, we take these to be the allowed energy levels of the electron. Energy is plotted vertically with the lowest or ground state at the bottom and with excited states above. Given the energies of the lines in an atomic spectrum, it is possible (although sometimes very difficult) to determine the energy levels of an atom. Energy-level diagrams are used for many systems, including molecules and nuclei. A theory of the atom or any other system must predict its energies based on the physics of the system.

Bohr was clever enough to find a way to calculate the electron orbital energies in hydrogen. This was an important first step that has been improved upon, but it is well worth repeating here, because it does correctly describe many characteristics of hydrogen. Assuming circular orbits, Bohr proposed that the angular momentum \( L \) of an electron in its orbit is quantized, that is, it has only specific, discrete values. The value for \( L \) is given by the formula

\[
L = m_e r_n v = n\frac{\hbar}{2\pi} \quad (n = 1, 2, 3, \ldots),
\]

where \( L \) is the angular momentum, \( m_e \) is the electron's mass, \( r_n \) is the radius of the \( n \)th orbit, and \( \hbar \) is Planck's constant. Note that angular momentum is \( L = I\omega \). For a small object at a radius \( r \), \( I = mr^2 \) and \( \omega = v/r \), so that \( L = (mv^2/r)v/r = mvr \). Quantization says that this value of \( mvr \) can only be equal to \( \hbar/2, 2\hbar/2, 3\hbar/2 \), etc. At the time, Bohr himself did not know why angular momentum should be quantized, but using this assumption he was able to calculate the energies in the hydrogen spectrum, something no one else had done at the time.

From Bohr's assumptions, we will now derive a number of important properties of the hydrogen atom from the classical physics we have covered in the text. We start by noting the centripetal force causing the electron to follow a circular path is supplied by the Coulomb force. To be more general, we note that this analysis is valid for any single-electron atom. So, if a nucleus has \( Z \) protons (\( Z = 1 \) for hydrogen, 2 for helium, etc.) and only one electron, that atom is called a **hydrogen-like atom**. The spectra of hydrogen-like ions are similar to hydrogen, but shifted to higher energy by the
greater attractive force between the electron and nucleus. The magnitude of the centripetal force is \( m_e v^2 / r_n \), while the Coulomb force is \( k(Zq_e)(q_e) / r_n^2 \). The tacit assumption here is that the nucleus is more massive than the stationary electron, and the electron orbits about it. This is consistent with the planetary model of the atom. Equating these,

\[
k\frac{q_e^2}{r_n^2} = \frac{m_e v^2}{r_n} \quad \text{(Coulomb = centripetal)}.
\]  

Angular momentum quantization is stated in an earlier equation. We solve that equation for \( v \), substitute it into the above, and rearrange the expression to obtain the radius of the orbit. This yields:

\[
r_n = n^2 a_B, \quad \text{for allowed orbits (} n = 1, 2, 3, \ldots \text{)},
\]  

where \( a_B \) is defined to be the Bohr radius, since for the lowest orbit (\( n = 1 \)) and for hydrogen (\( Z = 1 \)), \( r_1 = a_B \). It is left for this chapter’s Problems and Exercises to show that the Bohr radius is

\[
a_B = \frac{\hbar^2}{4\pi^2 m_e k q_e^2} = 0.529 \times 10^{-10} \text{ m}.
\]  

These last two equations can be used to calculate the radii of the allowed (quantized) electron orbits in any hydrogen-like atom. It is impressive that the formula gives the correct size of hydrogen, which is measured experimentally to be very close to the Bohr radius. The earlier equation also tells us that the orbital radius is proportional to \( n^2 \), as illustrated in Figure 30.19.

Figure 30.19 The allowed electron orbits in hydrogen have the radii shown. These radii were first calculated by Bohr and are given by the equation \( r_n = n^2 a_B \). The lowest orbit has the experimentally verified diameter of a hydrogen atom.

To get the electron orbital energies, we start by noting that the electron energy is the sum of its kinetic and potential energy:

\[
E_n = KE + PE.
\]  

Kinetic energy is the familiar \( KE = (1/2) m_e v^2 \), assuming the electron is not moving at relativistic speeds. Potential energy for the electron is electrical, or \( PE = q_e V \), where \( V \) is the potential due to the nucleus, which looks like a point charge. The nucleus has a positive charge \( Zq_e \); thus, \( V = kZq_e / r_n \), recalling an earlier equation for the potential due to a point charge. Since the electron’s charge is negative, we see that \( PE = -kZq_e / r_n \). Entering the expressions for \( KE \) and \( PE \), we find

\[
E_n = \frac{1}{2} m_e v^2 - kZq_e^2 r_n.
\]  

Now we substitute \( r_n \) and \( v \) from earlier equations into the above expression for energy. Algebraic manipulation yields

\[
E_n = -\frac{Z^2}{n^2} E_0 (n = 1, 2, 3, \ldots)
\]  

for the orbital energies of hydrogen-like atoms. Here, \( E_0 \) is the ground-state energy (\( n = 1 \)) for hydrogen (\( Z = 1 \)) and is given by
Thus, for hydrogen,

\[ E_n = -\frac{13.6 \text{ eV}}{n^2} (n = 1, 2, 3, \ldots) \]  

\[(30.28)\]

Figure 30.20 shows an energy-level diagram for hydrogen that also illustrates how the various spectral series for hydrogen are related to transitions between energy levels.

Electron total energies are negative, since the electron is bound to the nucleus, analogous to being in a hole without enough kinetic energy to escape. As \( n \) approaches infinity, the total energy becomes zero. This corresponds to a free electron with no kinetic energy, since \( r_n \) gets very large for large \( n \), and the electric potential energy thus becomes zero. Thus, 13.6 eV is needed to ionize hydrogen (to go from \(-13.6 \text{ eV}\) to 0, or unbound), an experimentally verified number. Given more energy, the electron becomes unbound with some kinetic energy. For example, giving 15.0 eV to an electron in the ground state of hydrogen strips it from the atom and leaves it with 1.4 eV of kinetic energy.

Finally, let us consider the energy of a photon emitted in a downward transition, given by the equation to be

\[ \Delta E = hf = E_i - E_f. \]  

\[(30.29)\]

Substituting \( E_n = (−13.6 \text{ eV} / n^2) \), we see that

\[ hf = (13.6 \text{ eV}) \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right). \]  

\[(30.30)\]

Dividing both sides of this equation by \( hc \) gives an expression for \( 1 / \lambda \): 

\[ \frac{hf}{hc} = \frac{f}{c} = \frac{1}{\lambda} = (13.6 \text{ eV}) \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right). \]  

\[(30.31)\]

It can be shown that

\[ \left( \frac{13.6 \text{ eV}}{hc} \right) = \left( \frac{13.6 \text{ eV}}{6.626 \times 10^{-34} \text{ J} \cdot \text{s}} \right) \left( \frac{1.602 \times 10^{-19} \text{ J/eV}}{2.998 \times 10^8 \text{ m/s}} \right) = 1.097 \times 10^7 \text{ m}^{-1} = R \]  

\[(30.32)\]
is the Rydberg constant. Thus, we have used Bohr's assumptions to derive the formula first proposed by Balmer years earlier as a recipe to fit experimental data.

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

(30.33)

We see that Bohr's theory of the hydrogen atom answers the question as to why this previously known formula describes the hydrogen spectrum. It is because the energy levels are proportional to \( 1/n^2 \), where \( n \) is a non-negative integer. A downward transition releases energy, and so \( n_i \) must be greater than \( n_f \). The various series are those where the transitions end on a certain level. For the Lyman series, \( n_f = 1 \) — that is, all the transitions end in the ground state (see also Figure 30.20). For the Balmer series, \( n_f = 2 \), or all the transitions end in the first excited state; and so on. What was once a recipe is now based in physics, and something new is emerging—angular momentum is quantized.

Triumphs and Limits of the Bohr Theory

Bohr did what no one had been able to do before. Not only did he explain the spectrum of hydrogen, he correctly calculated the size of the atom from basic physics. Some of his ideas are broadly applicable. Electron orbital energies are quantized in all atoms and molecules. Angular momentum is quantized. The electrons do not spiral into the nucleus, as expected classically (accelerated charges radiate, so that the electron orbits classically would decay quickly, and the electrons would sit on the nucleus—matter would collapse). These are major triumphs.

But there are limits to Bohr's theory. It cannot be applied to multielectron atoms, even one as simple as a two-electron helium atom. Bohr's model is what we call semiclasical. The orbits are quantized (nonclassical) but are assumed to be simple circular paths (classical). As quantum mechanics was developed, it became clear that there are no well-defined orbits; rather, there are clouds of probability. Bohr's theory also did not explain that some spectral lines are doublets (split into two) when examined closely. We shall examine many of these aspects of quantum mechanics in more detail, but it should be kept in mind that Bohr did not fail. Rather, he made very important steps along the path to greater knowledge and laid the foundation for all of atomic physics that has since evolved.

PhET Explorations: Models of the Hydrogen Atom

How did scientists figure out the structure of atoms without looking at them? Try out different models by shooting light at the atom. Check how the prediction of the model matches the experimental results.

PhET Interactive Simulation

Figure 30.21 Models of the Hydrogen Atom (http://cnx.org/content/m42596/1.4/hydrogen-atom_en.jar)

30.4 X Rays: Atomic Origins and Applications

Each type of atom (or element) has its own characteristic electromagnetic spectrum. X rays lie at the high-frequency end of an atom’s spectrum and are characteristic of the atom as well. In this section, we explore characteristic x rays and some of their important applications.

We have previously discussed x rays as a part of the electromagnetic spectrum in Photon Energies and the Electromagnetic Spectrum. That module illustrated how an x-ray tube (a specialized CRT) produces x rays. Electrons emitted from a hot filament are accelerated with a high voltage, gaining significant kinetic energy and striking the anode.

There are two processes by which x rays are produced in the anode of an x-ray tube. In one process, the deceleration of electrons produces x rays, and these x rays are called bremsstrahlung, or braking radiation. The second process is atomic in nature and produces characteristic x rays, so called because they are characteristic of the anode material. The x-ray spectrum in Figure 30.22 is typical of what is produced by an x-ray tube, showing a broad curve of bremsstrahlung radiation with characteristic x-ray peaks on it.

\[
qV = hf_{max}
\]

Figure 30.22 X-ray spectrum obtained when energetic electrons strike a material, such as in the anode of a CRT. The smooth part of the spectrum is bremsstrahlung radiation, while the peaks are characteristic of the anode material. A different anode material would have characteristic x-ray peaks at different frequencies.

The spectrum in Figure 30.22 is collected over a period of time in which many electrons strike the anode, with a variety of possible outcomes for each hit. The broad range of x-ray energies in the bremsstrahlung radiation indicates that an incident electron's energy is not usually converted.
entirely into photon energy. The highest-energy x-ray produced is one for which all of the electron’s energy was converted to photon energy. Thus the accelerating voltage and the maximum x-ray energy are related by conservation of energy. Electric potential energy is converted to kinetic energy and then to photon energy, so that 

\[ E_{\text{max}} = hf_{\text{max}} = q_e V. \]

Units of electron volts are convenient. For example, a 100-kV accelerating voltage produces x-ray photons with a maximum energy of 100 keV.

Some electrons excite atoms in the anode. Part of the energy that they deposit by collision with an atom results in one or more of the atom’s inner electrons being knocked into a higher orbit or the atom being ionized. When the anode’s atoms de-excite, they emit characteristic electromagnetic radiation. The most energetic of these are produced when an inner-shell vacancy is filled—that is, when an \( n = 1 \) or \( n = 2 \) shell electron has been excited to a higher level, and another electron falls into the vacant spot. A characteristic x ray (see Photon Energies and the Electromagnetic Spectrum) is electromagnetic (EM) radiation emitted by an atom when an inner-shell vacancy is filled. Figure 30.23 shows a representative energy-level diagram that illustrates the labeling of characteristic x rays. X rays created when an electron falls into an \( n = 1 \) shell vacancy are called \( K_{\alpha} \) when they come from the next higher level; that is, an \( n = 2 \) to \( n = 1 \) transition. The labels \( K, L, M,... \) come from the older alphabetical labeling of shells starting with \( K \) rather than using the principal quantum numbers 1, 2, 3, .... A more energetic \( K_{\beta} \) x ray is produced when an electron falls into an \( n = 1 \) shell vacancy from the \( n = 3 \) shell; that is, an \( n = 3 \) to \( n = 1 \) transition. Similarly, when an electron falls into the \( n = 2 \) shell from the \( n = 3 \) shell, an \( L_{\alpha} \) x ray is created. The energies of these x rays depend on the energies of electron states in the particular atom and, thus, are characteristic of that element: every element has its own set of x-ray energies. This property can be used to identify elements, for example, to find trace (small) amounts of an element in an environmental or biological sample.

Figure 30.23 A characteristic x ray is emitted when an electron fills an inner-shell vacancy, as shown for several transitions in this approximate energy level diagram for a multiple-electron atom. Characteristic x rays are labeled according to the shell that had the vacancy and the shell from which the electron came. A \( K_{\alpha} \) x ray, for example, is produced when an electron coming from the \( n = 2 \) shell fills the \( n = 1 \) shell vacancy.

Example 30.2 Characteristic X-Ray Energy

Calculate the approximate energy of a \( K_{\alpha} \) x ray from a tungsten anode in an x-ray tube.

**Strategy**

How do we calculate energies in a multiple-electron atom? In the case of characteristic x rays, the following approximate calculation is reasonable. Characteristic x rays are produced when an inner-shell vacancy is filled. Inner-shell electrons are nearer to the nucleus than others in an atom and thus feel little net effect from the others. This is similar to what happens inside a charged conductor, where its excess charge is distributed over the surface so that it produces no electric field inside. It is reasonable to assume the inner-shell electrons have hydrogen-like energies, as given by 

\[ E_n = -\frac{Z^2}{n^2}E_0 \]

(\( n = 1, 2, 3, ... \)). As noted, a \( K_{\alpha} \) x ray is produced by an \( n = 2 \) to \( n = 1 \) transition. Since there are two electrons in a filled \( K \) shell, a vacancy would leave one electron, so the effective charge would be \( Z - 1 \) rather than \( Z \). For tungsten, \( Z = 74 \), so that the effective charge is 73.

**Solution**

\[ E_n = -\frac{Z^2}{n^2}E_0 \]

(\( n = 1, 2, 3, ... \)) gives the orbital energies for hydrogen-like atoms to be 

\[ E_n = -(Z^2/n^2)E_0, \]

where \( E_0 = 13.6 \text{ eV} \). As noted, the effective \( Z \) is 73. Now the \( K_{\alpha} \) x-ray energy is given by

\[ E_{K_{\alpha}} = \Delta E = E_1 - E_1 = E_2 - E_1, \]

(30.34)

where

\[ E_1 = -\frac{Z^2}{1^2}E_0 = -\frac{73^2}{1}(13.6 \text{ eV}) = -72.5 \text{ keV} \]

(30.35)
\[ E_2 = -\frac{Z^2}{2}E_0 = -\frac{23^2}{4}(13.6 \text{ eV}) = -18.1 \text{ keV}. \]  \hfill (30.36)

Thus,
\[ E_{K\alpha} = -18.1 \text{ keV} \left( -72.5 \text{ keV} \right) = 54.4 \text{ keV}. \]  \hfill (30.37)

**Discussion**

This large photon energy is typical of characteristic x rays from heavy elements. It is large compared with other atomic emissions because it is produced when an inner-shell vacancy is filled, and inner-shell electrons are tightly bound. Characteristic x ray energies become progressively larger for heavier elements because their energy increases approximately as \( Z^2 \). Significant accelerating voltage is needed to create these inner-shell vacancies. In the case of tungsten, at least 72.5 kV is needed, because other shells are filled and you cannot simply bump one electron to a higher filled shell. Tungsten is a common anode material in x-ray tubes; so much of the energy of the impinging electrons is absorbed, raising its temperature, that a high-melting-point material like tungsten is required.

**Medical and Other Diagnostic Uses of X-rays**

All of us can identify diagnostic uses of x-ray photons. Among these are the universal dental and medical x rays that have become an essential part of medical diagnostics. (See Figure 30.25 and Figure 30.26.) X rays are also used to inspect our luggage at airports, as shown in Figure 30.24, and for early detection of cracks in crucial aircraft components. An x ray is not only a noun meaning high-energy photon, it also is an image produced by x rays, and it has been made into a familiar verb—to be x-rayed.

**Medical and Other Diagnostic Uses of X-rays**

All of us can identify diagnostic uses of x-ray photons. Among these are the universal dental and medical x rays that have become an essential part of medical diagnostics. (See Figure 30.25 and Figure 30.26.) X rays are also used to inspect our luggage at airports, as shown in Figure 30.24, and for early detection of cracks in crucial aircraft components. An x ray is not only a noun meaning high-energy photon, it also is an image produced by x rays, and it has been made into a familiar verb—to be x-rayed.

**Medical and Other Diagnostic Uses of X-rays**

All of us can identify diagnostic uses of x-ray photons. Among these are the universal dental and medical x rays that have become an essential part of medical diagnostics. (See Figure 30.25 and Figure 30.26.) X rays are also used to inspect our luggage at airports, as shown in Figure 30.24, and for early detection of cracks in crucial aircraft components. An x ray is not only a noun meaning high-energy photon, it also is an image produced by x rays, and it has been made into a familiar verb—to be x-rayed.

The most common x-ray images are simple shadows. Since x-ray photons have high energies, they penetrate materials that are opaque to visible light. The more energy an x-ray photon has, the more material it will penetrate. So an x-ray tube may be operated at 50.0 kV for a chest x ray, whereas it may need to be operated at 100 kV to examine a broken leg in a cast. The depth of penetration is related to the density of the material as well as to the energy of the photon. The denser the material, the fewer x-ray photons get through and the darker the shadow. Thus x rays excel at detecting breaks in bones and in imaging other physiological structures, such as some tumors, that differ in density from surrounding material. Because of their high photon energy, x rays produce significant ionization in materials and damage cells in biological organisms. Modern uses minimize exposure to the patient and eliminate exposure to others. Biological effects of x rays will be explored in the next chapter along with other types of ionizing radiation such as those produced by nuclei.
As the x-ray energy increases, the Compton effect (see Photon Momentum) becomes more important in the attenuation of the x rays. Here, the x-ray scatters from an outer electron shell of the atom, giving the ejected electron some kinetic energy while losing energy itself. The probability for attenuation of the x rays depends upon the number of electrons present (the material’s density) as well as the thickness of the material. Chemical composition of the medium, as characterized by its atomic number $Z$, is not important here. Low-energy x rays provide better contrast (sharper shadows); however, since atoms are on the order of 0.1 nm in size, x rays can be used to detect the location, shape, and size of atoms and molecules. The process is called x-ray diffraction, because it involves the diffraction and interference of x rays to produce patterns that can be analyzed for information about the structures that scattered the x rays. Perhaps the most famous example of x-ray diffraction is the discovery of the double-helix structure of DNA in 1953 by an international team of scientists working at the Cavendish Laboratory—American James Watson, Englishman Francis Crick, and New Zealand–born Maurice Wilkins. Using x-ray diffraction data produced by Rosalind Franklin, they were the first to discern the structure of DNA that is so crucial to life. For this, Watson, Crick, and Wilkins were awarded the 1962 Nobel Prize in Physiology or Medicine. There is much debate and controversy over the issue that Rosalind Franklin was not included in the prize.

Figure 30.29 shows a diffraction pattern produced by the scattering of x rays from a crystal. This process is known as x-ray crystallography because of the information it can yield about crystal structure, and it was the type of data Rosalind Franklin supplied to Watson and Crick for DNA. Not only do x rays confirm the size and shape of atoms, they give information on the atomic arrangements in materials. For example, current research in high-temperature superconductors involves complex materials whose lattice arrangements are crucial to obtaining a superconducting material. These can be studied using x-ray crystallography.

X-Ray Diffraction and Crystallography

Since x-ray photons are very energetic, they have relatively short wavelengths. For example, the 54.4-keV $K_{\alpha}$ x-ray of Example 30.2 has a wavelength $\lambda = \frac{hc}{E} = 0.0228$ nm. Thus, typical x-ray photons act like rays when they encounter macroscopic objects, like teeth, and produce sharp shadows; however, since atoms are on the order of 0.1 nm in size, x rays can be used to detect the location, shape, and size of atoms and molecules. The process is called x-ray diffraction, because it involves the diffraction and interference of x rays to produce patterns that can be analyzed for information about the structures that scattered the x rays. Perhaps the most famous example of x-ray diffraction is the discovery of the double-helix structure of DNA in 1953 by an international team of scientists working at the Cavendish Laboratory—American James Watson, Englishman Francis Crick, and New Zealand–born Maurice Wilkins. Using x-ray diffraction data produced by Rosalind Franklin, they were the first to discern the structure of DNA that is so crucial to life. For this, Watson, Crick, and Wilkins were awarded the 1962 Nobel Prize in Physiology or Medicine. There is much debate and controversy over the issue that Rosalind Franklin was not included in the prize.

Figure 30.29 shows a diffraction pattern produced by the scattering of x rays from a crystal. This process is known as x-ray crystallography because of the information it can yield about crystal structure, and it was the type of data Rosalind Franklin supplied to Watson and Crick for DNA. Not only do x rays confirm the size and shape of atoms, they give information on the atomic arrangements in materials. For example, current research in high-temperature superconductors involves complex materials whose lattice arrangements are crucial to obtaining a superconducting material. These can be studied using x-ray crystallography.
Historically, the scattering of x rays from crystals was used to prove that x rays are energetic EM waves. This was suspected from the time of the discovery of x rays in 1895, but it was not until 1912 that the German Max von Laue (1879–1960) convinced two of his colleagues to scatter x rays from crystals. If a diffraction pattern is obtained, he reasoned, then the x rays must be waves, and their wavelength could be determined. (The spacing of atoms in various crystals was reasonably well known at the time, based on good values for Avogadro’s number.) The experiments were convincing, and the 1914 Nobel Prize in Physics was given to von Laue for his suggestion leading to the proof that x rays are EM waves. In 1915, the unique father-and-son team of Sir William Henry Bragg and his son Sir William Lawrence Bragg were awarded a joint Nobel Prize for inventing the x-ray spectrometer and the then-new science of x-ray analysis. The elder Bragg had migrated to Australia from England just after graduating in mathematics. He learned physics and chemistry during his career at the University of Adelaide. The younger Bragg was born in Adelaide but went back to the Cavendish Laboratories in England to a career in x-ray and neutron crystallography; he provided support for Watson, Crick, and Wilkins for their work on unraveling the mysteries of DNA and to Max Perutz for his 1962 Nobel Prize-winning work on the structure of hemoglobin. Here again, we witness the enabling nature of physics—establishing instruments and designing experiments as well as solving mysteries in the biomedical sciences.

Certain other uses for x rays will be studied in later chapters. X rays are useful in the treatment of cancer because of the inhibiting effect they have on cell reproduction. X rays observed coming from outer space are useful in determining the nature of their sources, such as neutron stars and possibly black holes. Created in nuclear bomb explosions, x rays can also be used to detect clandestine atmospheric tests of these weapons. X rays can cause excitations of atoms, which then fluoresce (emitting characteristic EM radiation), making x-ray-induced fluorescence a valuable analytical tool in a range of fields from art to archaeology.

### 30.5 Applications of Atomic Excitations and De-Excitations

Many properties of matter and phenomena in nature are directly related to atomic energy levels and their associated excitations and de-exitations. The color of a rose, the output of a laser, and the transparency of air are but a few examples. (See Figure 30.30.) While it may not appear that glow-in-the-dark pajamas and lasers have much in common, they are in fact different applications of similar atomic de-excitations.

The color of a material is due to the ability of its atoms to absorb certain wavelengths while reflecting or reemitting others. A simple red material, for example a tomato, absorbs all visible wavelengths except red. This is because the atoms of its hydrocarbon pigment (lycopene) have levels separated by a variety of energies corresponding to all visible photon energies except red. Air is another interesting example. It is transparent to visible light, because there are few energy levels that visible photons can excite in air molecules and atoms. Visible light, thus, cannot be absorbed. Furthermore, visible light is only weakly scattered by air, because visible wavelengths are so much greater than the sizes of the air molecules and atoms. Light must pass through kilometers of air to scatter enough to cause red sunsets and blue skies.

#### Fluorescence and Phosphorescence

The ability of a material to emit various wavelengths of light is similarly related to its atomic energy levels. Figure 30.31 shows a scorpion illuminated by a UV lamp, sometimes called a black light. Some rocks also glow in black light, the particular colors being a function of the rock’s mineral composition. Black lights are also used to make certain posters glow.
In the fluorescence process, an atom is excited to a level several steps above its ground state by the absorption of a relatively high-energy UV photon. This is called **atomic excitation**. Once it is excited, the atom can de-excite in several ways, one of which is to re-emit a photon of the same energy as excited it, a single step back to the ground state. This is called **atomic de-excitation**. All other paths of de-excitation involve smaller steps, in which lower-energy (longer wavelength) photons are emitted. Some of these may be in the visible range, such as for the scorpion in Figure 30.31. **Fluorescence** is defined to be any process in which an atom or molecule, excited by a photon of a given energy, and de-excites by emission of a lower-energy photon.

Fluorescence can be induced by many types of energy input. Fluorescent paint, dyes, and even soap residues in clothes make colors seem brighter in sunlight by converting some UV into visible light. X rays can induce fluorescence, as is done in x-ray fluoroscopy to make brighter visible images. Electric discharges can induce fluorescence, as in so-called neon lights and in gas-discharge tubes that produce atomic and molecular spectra. Common fluorescent lights use an electric discharge in mercury vapor to cause atomic emissions from mercury atoms. The inside of a fluorescent light is coated with a fluorescent material that emits visible light over a broad spectrum of wavelengths. By choosing an appropriate coating, fluorescent lights can be made more like sunlight or like the reddish glow of candlelight, depending on needs. Fluorescent lights are more efficient in converting electrical energy into visible light than incandescent filaments (about four times as efficient), the blackbody radiation of which is primarily in the infrared due to temperature limitations.

This atom is excited to one of its higher levels by absorbing a UV photon. It can de-excite in a single step, re-emitting a photon of the same energy, or in several steps. The process is called fluorescence if the atom de-excites in smaller steps, emitting energy different from that which excited it. Fluorescence can be induced by a variety of energy inputs, such as UV, x-rays, and electrical discharge.

The spectacular Waitomo caves on North Island in New Zealand provide a natural habitat for glow-worms. The glow-worms hang up to 70 silk threads of about 30 or 40 cm each to trap prey that fly towards them in the dark. The fluorescence process is very efficient, with nearly 100% of the energy input turning into light. (In comparison, fluorescent lights are about 20% efficient.) Fluorescence has many uses in biology and medicine. It is commonly used to label and follow a molecule within a cell. Such tagging allows one to study the structure of DNA and proteins. Fluorescent dyes and antibodies are usually used to tag the molecules, which are then illuminated with UV light and their emission of visible light is observed. Since the fluorescence of each element is characteristic, identification of elements within a sample can be done this way.

Figure 30.32 shows a commonly used fluorescent dye called fluorescein. Below that, Figure 30.33 reveals the diffusion of a fluorescent dye in water by observing it under UV light.

**Figure 30.31** Objects glow in the visible spectrum when illuminated by an ultraviolet (black) light. Emissions are characteristic of the mineral involved, since they are related to its energy levels. In the case of scorpions, proteins near the surface of their skin give off the characteristic blue glow. This is a colorful example of fluorescence in which excitation is induced by UV radiation while de-excitation occurs in the form of visible light. (credit: Ken Bosma, Flickr)

**Figure 30.32** Fluorescein, shown here in powder form, is used to dye laboratory samples. (credit: Benjah-bmm27, Wikimedia Commons)

**Figure 30.33** Here, fluorescent powder is added to a beaker of water. The mixture gives off a bright glow under ultraviolet light. (credit: Bricksnite, Wikimedia Commons)
Nano-Crystals

Recently, a new class of fluorescent materials has appeared—“nano-crystals.” These are single-crystal molecules less than 100 nm in size. The smallest of these are called “quantum dots.” These semiconductor indicators are very small (2–6 nm) and provide improved brightness. They also have the advantage that all colors can be excited with the same incident wavelength. They are brighter and more stable than organic dyes and have a longer lifetime than conventional phosphors. They have become an excellent tool for long-term studies of cells, including migration and morphology. (Figure 30.34.)

Figure 30.34 Microscopic image of chicken cells using nano-crystals of a fluorescent dye. Cell nuclei exhibit blue fluorescence while neurofilaments exhibit green. (credit: Weerapong Prasongschean, Wikimedia Commons)

Once excited, an atom or molecule will usually spontaneously de-excite quickly. (The electrons raised to higher levels are attracted to lower ones by the positive charge of the nucleus.) Spontaneous de-excitation has a very short mean lifetime of typically about $10^{-8}$ s. However, some levels have significantly longer lifetimes, ranging up to milliseconds to minutes or even hours. These energy levels are inhibited and are slow in de-exciting because their quantum numbers differ greatly from those of available lower levels. Although these level lifetimes are short in human terms, they are many orders of magnitude longer than is typical and, thus, are said to be metastable, meaning relatively stable. Phosphorescence is the de-excitation of a metastable state. Glow-in-the-dark materials, such as luminous dials on some watches and clocks and on children’s toys and pajamas, are made of phosphorescent substances. Visible light excites the atoms or molecules to metastable states that decay slowly, releasing the stored excitation energy partially as visible light. In some ceramics, atomic excitation energy can be frozen in after the ceramic has cooled from its firing. It is very slowly released, but the ceramic can be induced to phosphoresce by heating—a process called “thermoluminescence.” Since the release is slow, thermoluminescence can be used to date antiquities. The less light emitted, the older the ceramic. (See Figure 30.35.)

Figure 30.35 Atoms frozen in an excited state when this Chinese ceramic figure was fired can be stimulated to de-excite and emit EM radiation by heating a sample of the ceramic—a process called thermoluminescence. Since the states slowly de-excite over centuries, the amount of thermoluminescence decreases with age, making it possible to use this effect to date and authenticate antiquities. This figure dates from the 11th century. (credit: Vassil, Wikimedia Commons)

Lasers

Lasers today are commonplace. Lasers are used to read bar codes at stores and in libraries, laser shows are staged for entertainment, laser printers produce high-quality images at relatively low cost, and lasers send prodigious numbers of telephone messages through optical fibers. Among other things, lasers are also employed in surveying, weapons guidance, tumor eradication, retinal welding, and for reading music CDs and computer CD-ROMs.

Why do lasers have so many varied applications? The answer is that lasers produce single-wavelength EM radiation that is also very coherent—that is, the emitted photons are in phase. Laser output can, thus, be more precisely manipulated than incoherent mixed-wavelength EM radiation from other sources. The reason laser output is so pure and coherent is based on how it is produced, which in turn depends on a metastable state in the lasing material. Suppose a material had the energy levels shown in Figure 30.36. When energy is put into a large collection of these atoms, electrons are raised to all possible levels. Most return to the ground state in less than about $10^{-8}$ s, but those in the metastable state linger. This includes
those electrons originally excited to the metastable state and those that fell into it from above. It is possible to get a majority of the atoms into the metastable state, a condition called a population inversion.

Figure 30.36 (a) Energy-level diagram for an atom showing the first few states, one of which is metastable. (b) Massive energy input excites atoms to a variety of states. (c) Most states decay quickly, leaving electrons only in the metastable and ground state. If a majority of electrons are in the metastable state, a population inversion has been achieved.

Once a population inversion is achieved, a very interesting thing can happen, as shown in Figure 30.37. An electron spontaneously falls from the metastable state, emitting a photon. This photon finds another atom in the metastable state and stimulates it to decay, emitting a second photon of the same wavelength and in phase with the first, and so on. Stimulated emission is the emission of electromagnetic radiation in the form of photons of a given frequency, triggered by photons of the same frequency. For example, an excited atom, with an electron in an energy orbit higher than normal, releases a photon of a specific frequency when the electron drops back to a lower energy orbit. If this photon then strikes another electron in the same high-energy orbit in another atom, another photon of the same frequency is released. The emitted photons and the triggering photons are always in phase, have the same polarization, and travel in the same direction. The probability of absorption of a photon is the same as the probability of stimulated emission, and so a majority of atoms must be in the metastable state to produce energy. Einstein (again Einstein, and back in 1917!) was one of the important contributors to the understanding of stimulated emission of radiation. Among other things, Einstein was the first to realize that stimulated emission and absorption are equally probable. The laser acts as a temporary energy storage device that subsequently produces a massive energy output of single-wavelength, in-phase photons.

Figure 30.37 One atom in the metastable state spontaneously decays to a lower level, producing a photon that goes on to stimulate another atom to de-excite. The second photon has exactly the same energy and wavelength as the first and is in phase with it. Both go on to stimulate the emission of other photons. A population inversion is necessary for there to be a net production rather than a net absorption of the photons.

The name laser is an acronym for light amplification by stimulated emission of radiation, the process just described. The process was proposed and developed following the advances in quantum physics. A joint Nobel Prize was awarded in 1964 to American Charles Townes (1915–), and Nikolay Basov (1922–2001) and Aleksandr Prokhorov (1916–2002), from the Soviet Union, for the development of lasers. The Nobel Prize in 1981 went to Arthur Schawlow (1921-1999) for pioneering laser applications. The original devices were called masers, because they produced microwaves. The first working laser was created in 1960 at Hughes Research labs (CA) by T. Maiman. It used a pulsed high-powered flash lamp and a ruby rod to produce red light. Today the name laser is used for all such devices developed to produce a variety of wavelengths, including microwave, infrared, visible, and ultraviolet radiation. Figure 30.38 shows how a laser can be constructed to enhance the stimulated emission of radiation. Energy input can be from a flash tube, electrical discharge, or other sources, in a process sometimes called optical pumping. A large percentage of the original pumping energy is dissipated in other forms, but a population inversion must be achieved. Mirrors can be used to enhance stimulated emission by...
multiple passes of the radiation back and forth through the lasing material. One of the mirrors is semitransparent to allow some of the light to pass through. The laser output from a laser is a mere 1% of the light passing back and forth in a laser.

Figure 30.38 Typical laser construction has a method of pumping energy into the lasing material to produce a population inversion. (a) Spontaneous emission begins with some photons escaping and others stimulating further emissions. (b) and (c) Mirrors are used to enhance the probability of stimulated emission by passing photons through the material several times.

Lasers are constructed from many types of lasing materials, including gases, liquids, solids, and semiconductors. But all lasers are based on the existence of a metastable state or a phosphorescent material. Some lasers produce continuous output; others are pulsed in bursts as brief as $10^{-14}$ s. Some laser outputs are fantastically powerful—some greater than $10^{12}$ W—but the more common, everyday lasers produce something on the order of $10^{-3}$ W. The helium-neon laser that produces a familiar red light is very common. Figure 30.39 shows the energy levels of helium and neon, a pair of noble gases that work well together. An electrical discharge is passed through a helium-neon gas mixture in which the number of atoms of helium is ten times that of neon. The first excited state of helium is metastable and, thus, stores energy. This energy is easily transferred by collision to neon atoms, because they have an excited state at nearly the same energy as that in helium. That state in neon is also metastable, and this is the one that produces the laser output. (The most likely transition is to the nearby state, producing 1.96 eV photons, which have a wavelength of 633 nm and appear red.) A population inversion can be produced in neon, because there are so many more helium atoms and these put energy into the neon. Helium-neon lasers often have continuous output, because the population inversion can be maintained even while lasing occurs. Probably the most common lasers in use today, including the common laser pointer, are semiconductor or diode lasers, made of silicon. Here, energy is pumped into the material by passing a current in the device to excite the electrons. Special coatings on the ends and fine cleavings of the semiconductor material allow light to bounce back and forth and a tiny fraction to emerge as laser light. Diode lasers can usually run continually and produce outputs in the milliwatt range.

Figure 30.39 Energy levels in helium and neon. In the common helium-neon laser, an electrical discharge pumps energy into the metastable states of both atoms. The gas mixture has about ten times more helium atoms than neon atoms. Excited helium atoms easily de-excite by transferring energy to neon in a collision. A population inversion in neon is achieved, allowing lasing by the neon to occur.
There are many medical applications of lasers. Lasers have the advantage that they can be focused to a small spot. They also have a well-defined wavelength. Many types of lasers are available today that provide wavelengths from the ultraviolet to the infrared. This is important, as one needs to be able to select a wavelength that will be preferentially absorbed by the material of interest. Objects appear a certain color because they absorb all other visible colors incident upon them. What wavelengths are absorbed depends upon the energy spacing between electron orbitals in that molecule. Unlike the hydrogen atom, biological molecules are complex and have a variety of absorption wavelengths or lines. But these can be determined and used in the selection of a laser with the appropriate wavelength. Water is transparent to the visible spectrum but will absorb light in the UV and IR regions. Blood (hemoglobin) strongly reflects red but absorbs most strongly in the UV.

Laser surgery uses a wavelength that is strongly absorbed by the tissue it is focused upon. One example of a medical application of lasers is shown in Figure 30.40. A detached retina can result in total loss of vision. Burns made by a laser focused to a small spot on the retina form scar tissue that can hold the retina in place, salvaging the patient's vision. Other light sources cannot be focused as precisely as a laser due to refractive dispersion of different wavelengths. Similarly, laser surgery in the form of cutting or burning away tissue is made more accurate because laser output can be very precisely focused and is preferentially absorbed because of its single wavelength. Depending upon what part or layer of the retina needs repairing, the appropriate type of laser can be selected. For the repair of tears in the retina, a green argon laser is generally used. This light is absorbed well by tissues containing blood, so coagulation or "welding" of the tear can be done.

In dentistry, the use of lasers is rising. Lasers are most commonly used for surgery on the soft tissue of the mouth. They can be used to remove ulcers, stop bleeding, and reshape gum tissue. Their use in cutting into bones and teeth is not quite so common; here the erbium YAG (yttrium aluminum garnet) laser is used.

The massive combination of lasers shown in Figure 30.41 can be used to induce nuclear fusion, the energy source of the sun and hydrogen bombs. Since lasers can produce very high power in very brief pulses, they can be used to focus an enormous amount of energy on a small glass sphere containing fusion fuel. Not only does the incident energy increase the fuel temperature significantly so that fusion can occur, it also compresses the fuel to great density, enhancing the probability of fusion. The compression or implosion is caused by the momentum of the impinging laser photons.

Music CDs are now so common that vinyl records are quaint antiquities. CDs (and DVDs) store information digitally and have a much larger information-storage capacity than vinyl records. An entire encyclopedia can be stored on a single CD. Figure 30.42 illustrates how the information is stored and read from the CD. Pits made in the CD by a laser can be tiny and very accurately spaced to record digital information. These are read by having an inexpensive solid-state infrared laser beam scatter from pits as the CD spins, revealing their digital pattern and the information encoded upon them.
A CD has digital information stored in the form of laser-created pits on its surface. These in turn can be read by detecting the laser light scattered from the pit. Large information capacity is possible because of the precision of the laser. Shorter-wavelength lasers enable greater storage capacity.

Holograms, such as those in Figure 30.43, are true three-dimensional images recorded on film by lasers. Holograms are used for amusement, decoration on novelty items and magazine covers, security on credit cards and driver’s licenses (a laser and other equipment is needed to reproduce them), and for serious three-dimensional information storage. You can see that a hologram is a true three-dimensional image, because objects change relative position in the image when viewed from different angles.

The name hologram means “entire picture” (from the Greek holo, as in holistic), because the image is three-dimensional. Holography is the process of producing holograms and, although they are recorded on photographic film, the process is quite different from normal photography. Holography uses light interference or wave optics, whereas normal photography uses geometric optics. Figure 30.44 shows one method of producing a hologram. Coherent light from a laser is split by a mirror, with part of the light illuminating the object. The remainder, called the reference beam, shines directly on a piece of film. Light scattered from the object interferes with the reference beam, producing constructive and destructive interference. As a result, the exposed film looks foggy, but close examination reveals a complicated interference pattern stored on it. Where the interference was constructive, the film (a negative actually) is darkened. Holography is sometimes called lensless photography, because it uses the wave characteristics of light as contrasted to normal photography, which uses geometric optics and so requires lenses.

Light falling on a hologram can form a three-dimensional image. The process is complicated in detail, but the basics can be understood as shown in Figure 30.45, in which a laser of the same type that exposed the film is now used to illuminate it. The myriad tiny exposed regions of the film are dark and block the light, while less exposed regions allow light to pass. The film thus acts much like a collection of diffraction gratings with various spacings. Light passing through the hologram is diffracted in various directions, producing both real and virtual images of the object used to expose
the film. The interference pattern is the same as that produced by the object. Moving your eye to various places in the interference pattern gives you different perspectives, just as looking directly at the object would. The image thus looks like the object and is three-dimensional like the object.

Figure 30.45 A transmission hologram is one that produces real and virtual images when a laser of the same type as that which exposed the hologram is passed through it. Diffraction from various parts of the film produces the same interference pattern as the object that was used to expose it.

The hologram illustrated in Figure 30.45 is a transmission hologram. Holograms that are viewed with reflected light, such as the white light holograms on credit cards, are reflection holograms and are more common. White light holograms often appear a little blurry with rainbow edges, because the diffraction patterns of various colors of light are at slightly different locations due to their different wavelengths. Further uses of holography include all types of 3-D information storage, such as of statues in museums and engineering studies of structures and 3-D images of human organs. Invented in the late 1940s by Dennis Gabor (1900–1970), who won the 1971 Nobel Prize in Physics for his work, holography became far more practical with the development of the laser. Since lasers produce coherent single-wavelength light, their interference patterns are more pronounced. The precision is so great that it is even possible to record numerous holograms on a single piece of film by just changing the angle of the film for each successive image. This is how the holograms that move as you walk by them are produced—a kind of lensless movie.

In a similar way, in the medical field, holograms have allowed complete 3-D holographic displays of objects from a stack of images. Storing these images for future use is relatively easy. With the use of an endoscope, high-resolution 3-D holographic images of internal organs and tissues can be made.

30.6 The Wave Nature of Matter Causes Quantization

After visiting some of the applications of different aspects of atomic physics, we now return to the basic theory that was built upon Bohr’s atom.

Einstein once said it was important to keep asking the questions we eventually teach children not to ask. Why is angular momentum quantized? You already know the answer. Electrons have wave-like properties, as de Broglie later proposed. They can exist only where they interfere constructively, and only certain orbits meet proper conditions, as we shall see in the next module.

Following Bohr’s initial work on the hydrogen atom, a decade was to pass before de Broglie proposed that matter has wave properties. The wave-like properties of matter were subsequently confirmed by observations of electron interference when scattered from crystals. Electrons can exist only in locations where they interfere constructively. How does this affect electrons in atomic orbits? When an electron is bound to an atom, its wavelength must fit into a small space, something like a standing wave on a string. (See Figure 30.46.) Allowed orbits are those orbits in which an electron constructively interferes with itself. Not all orbits produce constructive interference. Thus only certain orbits are allowed—the orbits are quantized.

Figure 30.46 (a) Waves on a string have a wavelength related to the length of the string, allowing them to interfere constructively. (b) If we imagine the string bent into a closed circle, we get a rough idea of how electrons in circular orbits can interfere constructively. (c) If the wavelength does not fit into the circumference, the electron interferes destructively; it cannot exist in such an orbit.

For a circular orbit, constructive interference occurs when the electron’s wavelength fits neatly into the circumference, so that wave crests always align with crests and wave troughs align with troughs, as shown in Figure 30.46 (b). More precisely, when an integral multiple of the electron’s wavelength equals the circumference of the orbit, constructive interference is obtained. In equation form, the condition for constructive interference and an allowed electron orbit is

\[ n\lambda_n = 2\pi r_n(n = 1, 2, 3 \ldots), \]  

where \( \lambda_n \) is the electron’s wavelength and \( r_n \) is the radius of that circular orbit. The de Broglie wavelength is \( \lambda = h/p = h/mv \), and so here \( \lambda = h/m_e v \). Substituting this into the previous condition for constructive interference produces an interesting result:

This content is available for free at http://cnx.org/content/col11406/1.7
\[
\frac{nh}{m_e v} = 2\pi r_n.
\]

Rearranging terms, and noting that \( L = mvr \) for a circular orbit, we obtain the quantization of angular momentum as the condition for allowed orbits:

\[
L = m_e vr_n = n\frac{\hbar}{2\pi}(n = 1, 2, 3 \ldots).
\]

This is what Bohr was forced to hypothesize as the rule for allowed orbits, as stated earlier. We now realize that it is the condition for constructive interference of an electron in a circular orbit. Figure 30.47 illustrates this for \( n = 3 \) and \( n = 4 \).

**Waves and Quantization**

The wave nature of matter is responsible for the quantization of energy levels in bound systems. Only those states where matter interferes constructively exist, or are “allowed.” Since there is a lowest orbit where this is possible in an atom, the electron cannot spiral into the nucleus. It cannot exist closer to or inside the nucleus. The wave nature of matter is what prevents matter from collapsing and gives atoms their sizes.

![Figure 30.47](image)

**Figure 30.47** The third and fourth allowed circular orbits have three and four wavelengths, respectively, in their circumferences.

Because of the wave character of matter, the idea of well-defined orbits gives way to a model in which there is a cloud of probability, consistent with Heisenberg’s uncertainty principle. Figure 30.48 shows how this applies to the ground state of hydrogen. If you try to follow the electron in some well-defined orbit using a probe that has a small enough wavelength to get some details, you will instead knock the electron out of its orbit. Each measurement of the electron’s position will find it to be in a definite location somewhere near the nucleus. Repeated measurements reveal a cloud of probability like that in the figure, with each speck the location determined by a single measurement. There is not a well-defined, circular-orbit type of distribution. Nature again proves to be different on a small scale than on a macroscopic scale.

![Figure 30.48](image)

**Figure 30.48** The ground state of a hydrogen atom has a probability cloud describing the position of its electron. The probability of finding the electron is proportional to the darkness of the cloud. The electron can be closer or farther than the Bohr radius, but it is very unlikely to be a great distance from the nucleus.

There are many examples in which the wave nature of matter causes quantization in bound systems such as the atom. Whenever a particle is confined or bound to a small space, its allowed wavelengths are those which fit into that space. For example, the particle in a box model describes a particle free to move in a small space surrounded by impenetrable barriers. This is true in blackbody radiators (atoms and molecules) as well as in atomic and molecular spectra. Various atoms and molecules will have different sets of electron orbits, depending on the size and complexity of the system. When a system is large, such as a grain of sand, the tiny particle waves in it can fit in so many ways that it becomes impossible to see that the allowed states are discrete. Thus the correspondence principle is satisfied. As systems become large, they gradually look less grainy, and quantization becomes less evident. Unbound systems (small or not), such as an electron freed from an atom, do not have quantized energies, since their wavelengths are not constrained to fit in a certain volume.
30.7 Patterns in Spectra Reveal More Quantization

High-resolution measurements of atomic and molecular spectra show that the spectral lines are even more complex than they first appear. In this section, we will see that this complexity has yielded important new information about electrons and their orbits in atoms.

In order to explore the substructure of atoms (and knowing that magnetic fields affect moving charges), the Dutch physicist Hendrik Lorentz (1853–1930) suggested that his student Pieter Zeeman (1865–1943) study how spectra might be affected by magnetic fields. What they found became known as the Zeeman effect, which involved spectral lines being split into two or more separate emission lines by an external magnetic field, as shown in Figure 30.50. For their discoveries, Zeeman and Lorentz shared the 1902 Nobel Prize in Physics.

Zeeman splitting is complex. Some lines split into three lines, some into five, and so on. But one general feature is that the amount the split lines are separated is proportional to the applied field strength, indicating an interaction with a moving charge. The splitting means that the quantized energy of an orbit is affected by an external magnetic field, causing the orbit to have several discrete energies instead of one. Even without an external magnetic field, very precise measurements showed that spectral lines are doublets (split into two), apparently by magnetic fields within the atom itself.

Bohr’s theory of circular orbits is useful for visualizing how an electron’s orbit is affected by a magnetic field. The circular orbit forms a current loop, which creates a magnetic field of its own, \( B_{\text{orb}} \) as seen in Figure 30.51. Note that the orbital magnetic field \( B_{\text{orb}} \) and the orbital angular momentum \( L_{\text{orb}} \) are along the same line. The external magnetic field and the orbital magnetic field interact; a torque is exerted to align them. A torque rotating a system through some angle does work so that there is energy associated with this interaction. Thus, orbits at different angles to the external magnetic field have different energies. What is remarkable is that the energies are quantized—the magnetic field splits the spectral lines into several discrete lines that have different energies. This means that only certain angles are allowed between the orbital angular momentum and the external field, as seen in Figure 30.52.
The approximate picture of an electron in a circular orbit illustrates how the current loop produces its own magnetic field, called $B_{\text{orb}}$. It also shows how $B_{\text{orb}}$ is along the same line as the orbital angular momentum $L_{\text{orb}}$.

Only certain angles are allowed between the orbital angular momentum and an external magnetic field. This is implied by the fact that the Zeeman effect splits spectral lines into several discrete lines. Each line is associated with an angle between the external magnetic field and magnetic fields due to electrons and their orbits.

We already know that the magnitude of angular momentum is quantized for electron orbits in atoms. The new insight is that the direction of the orbital angular momentum is also quantized. The fact that the orbital angular momentum can have only certain directions is called space quantization. Like many aspects of quantum mechanics, this quantization of direction is totally unexpected. On the macroscopic scale, orbital angular momentum, such as that of the moon around the earth, can have any magnitude and be in any direction.

Detailed treatment of space quantization began to explain some complexities of atomic spectra, but certain patterns seemed to be caused by something else. As mentioned, spectral lines are actually closely spaced doublets, a characteristic called fine structure, as shown in Figure 30.53. The doublet changes when a magnetic field is applied, implying that whatever causes the doublet interacts with a magnetic field. In 1925, Sem Goudsmit and George Uhlenbeck, two Dutch physicists, successfully argued that electrons have properties analogous to a macroscopic charge spinning on its axis. Electrons, in fact, have an internal or intrinsic angular momentum called intrinsic spin $S$. Since electrons are charged, their intrinsic spin creates an intrinsic magnetic field $B_{\text{int}}$, which interacts with their orbital magnetic field $B_{\text{orb}}$. Furthermore, electron intrinsic spin is quantized in magnitude and direction, analogous to the situation for orbital angular momentum. The spin of the electron can have only one magnitude, and its direction can be at only one of two angles relative to a magnetic field, as seen in Figure 30.54. We refer to this as spin up or spin down for the electron. Each spin direction has a different energy; hence, spectroscopic lines are split into two. Spectral doublets are now understood as being due to electron spin.
Figure 30.53 Fine structure. Upon close examination, spectral lines are doublets, even in the absence of an external magnetic field. The electron has an intrinsic magnetic field that interacts with its orbital magnetic field.

Figure 30.54 The intrinsic magnetic field $B_{\text{int}}$ of an electron is attributed to its spin, $S$, roughly pictured to be due to its charge spinning on its axis. This is only a crude model, since electrons seem to have no size. The spin and intrinsic magnetic field of the electron can make only one of two angles with another magnetic field, such as that created by the electron’s orbital motion. Space is quantized for spin as well as for orbital angular momentum.

These two new insights—that the direction of angular momentum, whether orbital or spin, is quantized, and that electrons have intrinsic spin—help to explain many of the complexities of atomic and molecular spectra. In magnetic resonance imaging, it is the way that the intrinsic magnetic field of hydrogen and biological atoms interact with an external field that underlies the diagnostic fundamentals.

### 30.8 Quantum Numbers and Rules

Physical characteristics that are quantized—such as energy, charge, and angular momentum—are of such importance that names and symbols are given to them. The values of quantized entities are expressed in terms of **quantum numbers**, and the rules governing them are of the utmost importance in determining what nature is and does. This section covers some of the more important quantum numbers and rules—all of which apply in chemistry, material science, and far beyond the realm of atomic physics, where they were first discovered. Once again, we see how physics makes discoveries which enable other fields to grow.

The energy states of bound systems are quantized, because the particle wavelength can fit into the bounds of the system in only certain ways. This was elaborated for the hydrogen atom, for which the allowed energies are expressed as $E_n \propto 1/n^2$, where $n = 1, 2, 3, \ldots$. We define $n$ to be the principal quantum number that labels the basic states of a system. The lowest-energy state has $n = 1$, the first excited state has $n = 2$, and so on. Thus the allowed values for the principal quantum number are

$$n = 1, 2, 3, \ldots \quad (30.41)$$

This is more than just a numbering scheme, since the energy of the system, such as the hydrogen atom, can be expressed as some function of $n$, as can other characteristics (such as the orbital radii of the hydrogen atom).

The fact that the magnitude of angular momentum is quantized was first recognized by Bohr in relation to the hydrogen atom; it is now known to be true in general. With the development of quantum mechanics, it was found that the magnitude of angular momentum $L$ can have only the values
where \( l \) is defined to be the angular momentum quantum number. The rule for \( l \) in atoms is given in the parentheses. Given \( n \), the value of \( l \) can be any integer from zero up to \( n - 1 \). For example, if \( n = 4 \), then \( l \) can be 0, 1, 2, or 3.

Note that for \( n = 1 \), \( l \) can only be zero. This means that the ground-state angular momentum for hydrogen is actually zero, not \( \hbar/2\pi \) as Bohr proposed. The picture of circular orbits is not valid, because there would be angular momentum for any circular orbit. A more valid picture is the cloud of probability shown for the ground state of hydrogen in Figure 30.48. The electron actually spends time in and near the nucleus. The reason the electron does not remain in the nucleus is related to Heisenberg’s uncertainty principle—the electron’s energy would have to be much too large to be confined to the small space of the nucleus. Now the first excited state of hydrogen has \( n = 2 \), so that \( l \) can be either 0 or 1, according to the rule in

\[
L = \sqrt{l(l+1)} \frac{\hbar}{2\pi} \quad (l = 0, 1, 2, ..., n - 1),
\]

(30.42)

It is much simpler to state \( l = 2 \).

As recognized in the Zeeman effect, the direction of angular momentum is quantized. We now know this is true in all circumstances. It is found that the component of angular momentum along one direction in space, usually called the \( z \)-axis, can have only certain values of \( L_z \). The direction in space must be related to something physical, such as the direction of the magnetic field at that location. This is an aspect of relativity. Direction has no meaning if there is nothing that varies with direction, as does magnetic force. The allowed values of \( L_z \) are

\[
L_z = m_l \frac{\hbar}{2\pi} \quad (m_l = -l, -l + 1, ..., -1, 0, 1, ..., l - 1, l),
\]

(30.44)

where \( L_z \) is the \( z \)-component of the angular momentum and \( m_l \) is the angular momentum projection quantum number. The rule in parentheses for the values of \( m_l \) is that it can range from \(-l\) to \( l\) in steps of one. For example, if \( l = 2 \), then \( m_l \) can have the five values \(-2, -1, 0, 1, \) and \( 2 \). Each \( m_l \) corresponds to a different energy in the presence of a magnetic field, so that they are related to the splitting of spectral lines into discrete parts, as discussed in the preceding section. If the \( z \)-component of angular momentum can have only certain values, then the angular momentum can have only certain directions, as illustrated in Figure 30.55.

**Example 30.3 What Are the Allowed Directions?**

Calculate the angles that the angular momentum vector \( \mathbf{L} \) can make with the \( z \)-axis for \( l = 1 \), as illustrated in Figure 30.55.

**Strategy**
**Figure 30.55** represents the vectors \( \mathbf{L} \) and \( \mathbf{L}_z \) as usual, with arrows proportional to their magnitudes and pointing in the correct directions. \( \mathbf{L} \) and \( \mathbf{L}_z \) form a right triangle, with \( \mathbf{L} \) being the hypotenuse and \( \mathbf{L}_z \) the adjacent side. This means that the ratio of \( \mathbf{L}_z \) to \( \mathbf{L} \) is the cosine of the angle of interest. We can find \( \mathbf{L} \) and \( \mathbf{L}_z \) using

\[
\mathbf{L} = \sqrt{l(l+1)} \frac{\hbar}{2\pi} \quad \text{and} \quad \mathbf{L}_z = m \frac{\hbar}{2\pi}.
\]

**Solution**

We are given \( l = 1 \), so that \( m_l \) can be +1, 0, or −1. Thus \( \mathbf{L} \) has the value given by

\[
\mathbf{L} = \sqrt{l(l+1)} \frac{\hbar}{2\pi} = 2 \frac{\hbar}{2\pi}.
\]

(30.45)

\( \mathbf{L}_z \) can have three values, given by

\[
\mathbf{L}_z = m \frac{\hbar}{2\pi} = \begin{cases} 
\frac{\hbar}{2\pi}, & m_l = +1 \\
0, & m_l = 0 \\
-\frac{\hbar}{2\pi}, & m_l = -1
\end{cases}
\]

(30.46)

As can be seen in **Figure 30.55**, \( \cos \theta = \mathbf{L}_z/\mathbf{L} \), and so for \( m_l = +1 \), we have

\[
\cos \theta_1 = \frac{L_z}{L} = \frac{\frac{\hbar}{2\pi}}{2} = \frac{1}{\sqrt{2}} = 0.707.
\]

Thus,

\[
\theta_1 = \cos^{-1}(0.707) = 45.0^\circ.
\]

(30.47)

Similarly, for \( m_l = 0 \), we find \( \cos \theta_2 = 0 \); thus,

\[
\theta_2 = \cos^{-1}(0) = 90.0^\circ.
\]

(30.49)

And for \( m_l = -1 \),

\[
\cos \theta_3 = \frac{L_z}{L} = \frac{-\frac{\hbar}{2\pi}}{2} = -\frac{1}{\sqrt{2}} = -0.707,
\]

so that

\[
\theta_3 = \cos^{-1}(-0.707) = 135.0^\circ.
\]

(30.51)

**Discussion**

The angles are consistent with the figure. Only the angle relative to the \( z \)-axis is quantized. \( \mathbf{L} \) can point in any direction as long as it makes the proper angle with the \( z \)-axis. Thus the angular momentum vectors lie on cones as illustrated. This behavior is not observed on the large scale.

To see how the correspondence principle holds here, consider that the smallest angle (\( \theta_1 \) in the example) is for the maximum value of \( m_l = 0 \), namely \( m_l = l \). For that smallest angle,

\[
\cos \theta = \frac{L_z}{L} = \frac{l}{\sqrt{l(l+1)}},
\]

(30.52)

which approaches 1 as \( l \) becomes very large. If \( \cos \theta = 1 \), then \( \theta = 0^\circ \). Furthermore, for large \( l \), there are many values of \( m_l \), so that all angles become possible as \( l \) gets very large.

---

**Intrinsic Spin Angular Momentum Is Quantized in Magnitude and Direction**

There are two more quantum numbers of immediate concern. Both were first discovered for electrons in conjunction with fine structure in atomic spectra. It is now well established that electrons and other fundamental particles have intrinsic spin, roughly analogous to a planet spinning on its axis. This spin is a fundamental characteristic of particles, and only one magnitude of intrinsic spin is allowed for a given type of particle. Intrinsic angular momentum is quantized independently of orbital angular momentum. Additionally, the direction of the spin is also quantized. It has been found that the magnitude of the intrinsic (internal) spin angular momentum, \( S \), of an electron is given by

\[
S = \sqrt{s(s+1)} \frac{\hbar}{2\pi} \quad (s = 1/2 \text{ for electrons}),
\]

(30.53)
where \( s \) is defined to be the spin quantum number. This is very similar to the quantization of \( L \) given in \( L = \sqrt{l(l + 1)} \frac{\hbar}{2\pi} \), except that the only value allowed for \( s \) for electrons is 1/2.

The direction of intrinsic spin is quantized, just as is the direction of orbital angular momentum. The direction of spin angular momentum along one direction in space, again called the \( z \)-axis, can have only the values

\[
S_z = m_s \frac{\hbar}{2\pi} \left( m_s = -\frac{1}{2}, +\frac{1}{2} \right)
\]

(30.54)

for electrons. \( S_z \) is the \( z \)-component of spin angular momentum and \( m_s \) is the spin projection quantum number. For electrons, \( s \) can only be 1/2, and \( m_s \) can be either +1/2 or –1/2. Spin projection \( m_s = +1/2 \) is referred to as spin up, whereas \( m_s = -1/2 \) is called spin down. These are illustrated in Figure 30.54.

### Intrinsic Spin

In later chapters, we will see that intrinsic spin is a characteristic of all subatomic particles. For some particles \( s \) is half-integral, whereas for others \( s \) is integral—there are crucial differences between half-integral spin particles and integral spin particles. Protons and neutrons, like electrons, have \( s = 1/2 \), whereas photons have \( s = 1 \), and other particles called pions have \( s = 0 \), and so on.

To summarize, the state of a system, such as the precise nature of an electron in an atom, is determined by its particular quantum numbers. These are expressed in the form \( (n, l, m_l, m_s) \)—see Table 30.1 For electrons in atoms, the principal quantum number can have the values \( n = 1, 2, 3, \ldots \). Once \( n \) is known, the values of the angular momentum quantum number are limited to \( l = 1, 2, 3, \ldots, n - 1 \). For a given value of \( l \), the angular momentum projection quantum number can have only the values \( m_l = -l, -l + 1, \ldots, -1, 0, 1, \ldots, l - 1, l \). Electron spin is independent of \( n, l, \) and \( m_l \), always having \( s = 1/2 \). The spin projection quantum number can have two values, \( m_s = 1/2 \) or \( -1/2 \).

### Table 30.1 Atomic Quantum Numbers

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Allowed values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal quantum number</td>
<td>( n )</td>
<td>1, 2, 3, ...</td>
</tr>
<tr>
<td>Angular momentum</td>
<td>( l )</td>
<td>0, 1, 2, ..., ( n - 1 )</td>
</tr>
<tr>
<td>Angular momentum projection</td>
<td>( m_l )</td>
<td>( -l, -l + 1, \ldots, -1, 0, 1, \ldots, l - 1, l ) (or 0, ( \pm 1, \pm 2, \ldots, \pm l ))</td>
</tr>
<tr>
<td>Spin</td>
<td>( s )</td>
<td>1/2(electrons)</td>
</tr>
<tr>
<td>Spin projection</td>
<td>( m_s )</td>
<td>( -1/2, +1/2 )</td>
</tr>
</tbody>
</table>

Figure 30.56 shows several hydrogen states corresponding to different sets of quantum numbers. Note that these clouds of probability are the locations of electrons as determined by making repeated measurements—each measurement finds the electron in a definite location, with a greater chance of finding the electron in some places rather than others. With repeated measurements, the pattern of probability shown in the figure emerges. The clouds of probability do not look like nor do they correspond to classical orbits. The uncertainty principle actually prevents us and nature from knowing how the electron gets from one place to another, and so an orbit really does not exist as such. Nature on a small scale is again much different from that on the large scale.

---

1. The spin quantum number \( s \) is usually not stated, since it is always 1/2 for electrons
Figure 30.56 Probability clouds for the electron in the ground state and several excited states of hydrogen. The nature of these states is determined by their sets of quantum numbers, here given as \((n, \ell, m)\). The ground state is \((0, 0, 0)\); one of the possibilities for the second excited state is \((3, 2, 1)\). The probability of finding the electron is indicated by the shade of color; the darker the coloring the greater the chance of finding the electron.

We will see that the quantum numbers discussed in this section are valid for a broad range of particles and other systems, such as nuclei. Some quantum numbers, such as intrinsic spin, are related to fundamental classifications of subatomic particles, and they obey laws that will give us further insight into the substructure of matter and its interactions.

PhET Explorations: Stern-Gerlach Experiment

The classic Stern-Gerlach Experiment shows that atoms have a property called spin. Spin is a kind of intrinsic angular momentum, which has no classical counterpart. When the z-component of the spin is measured, one always gets one of two values: spin up or spin down.

Figure 30.57 Stern-Gerlach Experiment (http://cnx.org/content/m42614/1.9/stern-gerlach_en.jar)

30.9 The Pauli Exclusion Principle

Multiple-Electron Atoms

All atoms except hydrogen are multiple-electron atoms. The physical and chemical properties of elements are directly related to the number of electrons a neutral atom has. The periodic table of the elements groups elements with similar properties into columns. This systematic organization is related to the number of electrons in a neutral atom, called the atomic number, \(Z\). We shall see in this section that the exclusion principle is key to the underlying explanations, and that it applies far beyond the realm of atomic physics.
In 1925, the Austrian physicist Wolfgang Pauli (see Figure 30.58) proposed the following rule: No two electrons can have the same set of quantum numbers. That is, no two electrons can be in the same state. This statement is known as the Pauli exclusion principle, because it excludes electrons from being in the same state. The Pauli exclusion principle is extremely powerful and very broadly applicable. It applies to any identical particles with half-integral intrinsic spin—that is, having \( s = 1/2, 3/2, \ldots \) Thus no two electrons can have the same set of quantum numbers.

### Pauli Exclusion Principle

No two electrons can have the same set of quantum numbers. That is, no two electrons can be in the same state.

---

**Figure 30.58** The Austrian physicist Wolfgang Pauli (1900–1958) played a major role in the development of quantum mechanics. He proposed the exclusion principle; hypothesized the existence of an important particle, called the neutrino, before it was directly observed; made fundamental contributions to several areas of theoretical physics; and influenced many students who went on to do important work of their own. (credit: Nobel Foundation, via Wikimedia Commons)

Let us examine how the exclusion principle applies to electrons in atoms. The quantum numbers involved were defined in [Quantum Numbers and Rules](Quantum%20Numbers%20and%20Rules) as \( n, l, m_l, s, \) and \( m_s \). Since \( s \) is always \( 1/2 \) for electrons, it is redundant to list \( s \), and so we omit it and specify the state of an electron by a set of four numbers \( (n, l, m_l, m_s) \). For example, the quantum numbers \( (2, 1, 0, -1/2) \) completely specify the state of an electron in an atom.

Since no two electrons can have the same set of quantum numbers, there are limits to how many of them can be in the same energy state. Note that \( n \) determines the energy state in the absence of a magnetic field. So we first choose \( n \), and then we see how many electrons can be in this energy state or energy level. Consider the \( n = 1 \) level, for example. The only value \( l \) can have is 0 (see [Table 30.1](Table%2030.1) for a list of possible values once \( n \) is known), and thus \( m_l \) can only be 0. The spin projection \( m_s \) can be either \( +1/2 \) or \( -1/2 \), and so there can be two electrons in the \( n = 1 \) state. One has quantum numbers \( (1, 0, 0, +1/2) \), and the other has \( (1, 0, 0, -1/2) \). **Figure 30.59** illustrates that there can be one or two electrons having \( n = 1 \), but not three.
The Pauli exclusion principle explains why some configurations of electrons are allowed while others are not. Since electrons cannot have the same set of quantum numbers, a maximum of two can be in the $n = 1$ level, and a third electron must reside in the higher-energy $n = 2$ level. If there are two electrons in the $n = 1$ level, their spins must be in opposite directions. (More precisely, their spin projections must differ.)

**Shells and Subshells**

Because of the Pauli exclusion principle, only hydrogen and helium can have all of their electrons in the $n = 1$ state. Lithium (see the periodic table) has three electrons, and so one must be in the $n = 2$ level. This leads to the concept of shells and shell filling. As we progress up in the number of electrons, we go from hydrogen to helium, lithium, beryllium, boron, and so on, and we see that there are limits to the number of electrons for each value of $n$. Higher values of the shell $n$ correspond to higher energies, and they can allow more electrons because of the various combinations of $l, m_l, \text{and } m_s$ that are possible. Each value of the principal quantum number $n$ thus corresponds to an atomic shell into which a limited number of electrons can go. Shells and the number of electrons in them determine the physical and chemical properties of atoms, since it is the outermost electrons that interact most with anything outside the atom.

The probability clouds of electrons with the lowest value of $l$ are closest to the nucleus and, thus, more tightly bound. Thus when shells fill, they start with $l = 0$, progress to $l = 1$, and so on. Each value of $l$ thus corresponds to a subshell.

The table given below lists symbols traditionally used to denote shells and subshells.

<table>
<thead>
<tr>
<th>Shell</th>
<th>Subshell</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>$s$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$p$</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>$d$</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>$f$</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>$g$</td>
</tr>
<tr>
<td>6$^{[2]}$</td>
<td>$h$</td>
<td></td>
</tr>
</tbody>
</table>

To denote shells and subshells, we write $nl$ with a number for $n$ and a letter for $l$. For example, an electron in the $n = 1$ state must have $l = 0$, and it is denoted as a $1s$ electron. Two electrons in the $n = 1$ state is denoted as $1s^2$. Another example is an electron in the $n = 2$ state with $l = 1$, written as $2p$. The case of three electrons with these quantum numbers is written $2p^3$. This notation, called spectroscopic notation, is generalized as shown in Figure 30.60.

2. It is unusual to deal with subshells having $l$ greater than 6, but when encountered, they continue to be labeled in alphabetical order.
Counting the number of possible combinations of quantum numbers allowed by the exclusion principle, we can determine how many electrons it takes to fill each subshell and shell.

**Example 30.4 How Many Electrons Can Be in This Shell?**

List all the possible sets of quantum numbers for the \( n = 2 \) shell, and determine the number of electrons that can be in the shell and each of its subshells.

**Strategy**

Given \( n = 2 \) for the shell, the rules for quantum numbers limit \( l \) to be 0 or 1. The shell therefore has two subshells, labeled \( 2s \) and \( 2p \). Since the lowest \( l \) subshell fills first, we start with the \( 2s \) subshell possibilities and then proceed with the \( 2p \) subshell.

**Solution**

It is convenient to list the possible quantum numbers in a table, as shown below.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \ell )</th>
<th>( m )</th>
<th>( m_s )</th>
<th><strong>Subshell</strong></th>
<th><strong>Total in subshell</strong></th>
<th><strong>Total in shell</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>+1/2</td>
<td>( 2s )</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>-1/2</td>
<td>( 2s )</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>+1/2</td>
<td>( 2p )</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-1/2</td>
<td>( 2p )</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>+1/2</td>
<td>( 2p )</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>-1/2</td>
<td>( 2p )</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1</td>
<td>+1/2</td>
<td>( 2p )</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1</td>
<td>-1/2</td>
<td>( 2p )</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

**Discussion**

It is laborious to make a table like this every time we want to know how many electrons can be in a shell or subshell. There exist general rules that are easy to apply, as we shall now see.

The number of electrons that can be in a subshell depends entirely on the value of \( l \). Once \( l \) is known, there are a fixed number of values of \( m_l \), each of which can have two values for \( m_s \). First, since \( m_l \) goes from \(-l \) to \( l \) in steps of 1, there are \( 2l + 1 \) possibilities. This number is multiplied by 2, since each electron can be spin up or spin down. Thus the **maximum number of electrons that can be in a subshell** is \( 2(2l + 1) \).

For example, the \( 2s \) subshell in **Example 30.4** has a maximum of 2 electrons in it, since \( 2(2(1) + 1) = 2(0 + 1) = 2 \) for this subshell. Similarly, the \( 2p \) subshell has a maximum of 6 electrons, since \( 2(2(1) + 1) = 2(2 + 1) = 6 \). For a shell, the maximum number is the sum of what can fit in the subshells. Some algebra shows that the **maximum number of electrons that can be in a shell** is \( 2n^2 \).

For example, for the first shell \( n = 1 \), and so \( 2n^2 = 2 \). We have already seen that only two electrons can be in the \( n = 1 \) shell. Similarly, for the second shell, \( n = 2 \), and so \( 2n^2 = 8 \). As found in **Example 30.4**, the total number of electrons in the \( n = 2 \) shell is 8.

**Example 30.5 Subshells and Totals for \( n = 3 \)**

How many subshells are in the \( n = 3 \) shell? Identify each subshell, calculate the maximum number of electrons that will fit into each, and verify that the total is \( 2n^2 \).

**Strategy**

Subshells are determined by the value of \( l \); thus, we first determine which values of \( l \) are allowed, and then we apply the equation “maximum number of electrons that can be in a subshell \( = 2(2l + 1) \)” to find the number of electrons in each subshell.

**Solution**
Since \( n = 3 \), we know that \( l \) can be 0, 1, or 2; thus, there are three possible subshells. In standard notation, they are labeled the 3s, 3p, and 3d subshells. We have already seen that 2 electrons can be in an s state, and 6 in a p state, but let us use the equation “maximum number of electrons that can be in a subshell = \( 2(2l + 1) \)” to calculate the maximum number in each:

\[
\begin{align*}
3s \text{ has } l &= 0; \text{ thus, } 2(2l + 1) = 2(0 + 1) = 2 \\
3p \text{ has } l &= 1; \text{ thus, } 2(2l + 1) = 2(2 + 1) = 6 \\
3d \text{ has } l &= 2; \text{ thus, } 2(2l + 1) = 2(4 + 1) = 10
\end{align*}
\]

Total = 18

(in the \( n = 3 \) shell)

The equation “maximum number of electrons that can be in a shell = \( 2n^2 \)” gives the maximum number in the \( n = 3 \) shell to be

\[
\text{Maximum number of electrons } = 2n^2 = 2(3)^2 = 2(9) = 18.
\]

Discussion

The total number of electrons in the three possible subshells is thus the same as the formula \( 2n^2 \). In standard (spectroscopic) notation, a filled \( n = 3 \) shell is denoted as \( 3s^2 3p^6 3d^{10} \). Shells do not fill in a simple manner. Before the \( n = 3 \) shell is completely filled, for example, we begin to find electrons in the \( n = 4 \) shell.

Shell Filling and the Periodic Table

Table 30.3 shows electron configurations for the first 20 elements in the periodic table, starting with hydrogen and its single electron and ending with calcium. The Pauli exclusion principle determines the maximum number of electrons allowed in each shell and subshell. But the order in which the shells and subshells are filled is complicated because of the large numbers of interactions between electrons.
<table>
<thead>
<tr>
<th>Element</th>
<th>Number of electrons (Z)</th>
<th>Ground state configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1</td>
<td>1s (^1)</td>
</tr>
<tr>
<td>He</td>
<td>2</td>
<td>1s (^2) 2s (^1)</td>
</tr>
<tr>
<td>Li</td>
<td>3</td>
<td>1s (^2) 2s (^1)</td>
</tr>
<tr>
<td>Be</td>
<td>4</td>
<td>2s (^2)</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>2s (^2) 2p (^1)</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>2s (^2) 2p (^2)</td>
</tr>
<tr>
<td>N</td>
<td>7</td>
<td>2s (^2) 2p (^3)</td>
</tr>
<tr>
<td>O</td>
<td>8</td>
<td>2s (^2) 2p (^4)</td>
</tr>
<tr>
<td>F</td>
<td>9</td>
<td>2s (^2) 2p (^5)</td>
</tr>
<tr>
<td>Ne</td>
<td>10</td>
<td>2s (^2) 2p (^6)</td>
</tr>
<tr>
<td>Na</td>
<td>11</td>
<td>2s (^2) 2p (^6) 3s (^1)</td>
</tr>
<tr>
<td>Mg</td>
<td>12</td>
<td>3s (^2)</td>
</tr>
<tr>
<td>Al</td>
<td>13</td>
<td>3s (^2) 3p (^1)</td>
</tr>
<tr>
<td>Si</td>
<td>14</td>
<td>3s (^2) 3p (^2)</td>
</tr>
<tr>
<td>P</td>
<td>15</td>
<td>3s (^2) 3p (^3)</td>
</tr>
<tr>
<td>S</td>
<td>16</td>
<td>3s (^2) 3p (^4)</td>
</tr>
<tr>
<td>Cl</td>
<td>17</td>
<td>3s (^2) 3p (^5)</td>
</tr>
<tr>
<td>Ar</td>
<td>18</td>
<td>3s (^2) 3p (^6)</td>
</tr>
<tr>
<td>K</td>
<td>19</td>
<td>3s (^2) 3p (^6) 4s (^1)</td>
</tr>
<tr>
<td>Ca</td>
<td>20</td>
<td>4s (^2)</td>
</tr>
</tbody>
</table>

Examining the above table, you can see that as the number of electrons in an atom increases from 1 in hydrogen to 2 in helium and so on, the lowest-energy shell gets filled first—that is, the \( n = 1 \) shell fills first, and then the \( n = 2 \) shell begins to fill. Within a shell, the subshells fill starting with the lowest \( l \), or with the \( s \) subshell, then the \( p \), and so on, usually until all subshells are filled. The first exception to this occurs for potassium, where the \( 4s \) subshell begins to fill before any electrons go into the \( 3d \) subshell. The next exception is not shown in Table 30.3; it occurs for rubidium, where the \( 5s \) subshell starts to fill before the \( 4d \) subshell. The reason for these exceptions is that \( l = 0 \) electrons have probability clouds that penetrate closer to the nucleus and, thus, are more tightly bound (lower in energy). Figure 30.62 shows the periodic table of the elements, through element 118. Of special interest are elements in the main groups, namely, those in the columns numbered 1, 2, 13, 14, 15, 16, 17, and 18.
The number of electrons in the outermost subshell determines the atom's chemical properties, since it is these electrons that are farthest from the nucleus and thus interact most with other atoms. If the outermost subshell can accept or give up an electron easily, then the atom will be highly reactive chemically. Each group in the periodic table is characterized by its outermost electron configuration. Perhaps the most familiar is Group 18 (Group VIII), the noble gases (helium, neon, argon, etc.). These gases are all characterized by a filled outer subshell that is particularly stable. This means that they have large ionization energies and do not readily give up an electron. Furthermore, if they were to accept an extra electron, it would be in a significantly higher level and thus loosely bound. Chemical reactions often involve sharing electrons. Noble gases can be forced into unstable chemical compounds only under high pressure and temperature.

Group 17 (Group VII) contains the halogens, such as fluorine, chlorine, iodine and bromine, each of which has one less electron than a neighboring noble gas. Each halogen has 5 \( p \) electrons (a \( p^5 \) configuration), while the \( p \) subshell can hold 6 electrons. This means the halogens have one vacancy in their outermost subshell. They thus readily accept an extra electron (it becomes tightly bound, closing the shell as in noble gases) and are highly reactive chemically. The halogens are also likely to form singly negative ions, such as \( \text{Cl}^- \), fitting an extra electron into the vacancy in the outer subshell. In contrast, alkali metals, such as sodium and potassium, all have a single \( s \) electron in their outermost subshell (an \( s^1 \) configuration) and are members of Group 1 (Group I). These elements easily give up their extra electron and are thus highly reactive chemically. As you might expect, they also tend to form singly positive ions, such as \( \text{Na}^+ \), by losing their loosely bound outermost electron. They are metals (conductors), because the loosely bound outer electron can move freely.

Of course, other groups are also of interest. Carbon, silicon, and germanium, for example, have similar chemistries and are in Group 4 (Group IV). Carbon, in particular, is extraordinary in its ability to form many types of bonds and to be part of long chains, such as inorganic molecules. The large group of what are called transitional elements is characterized by the filling of the \( d \) subshells and crossing of energy levels. Heavier groups, such as the lanthanide series, are more complex—their shells do not fill in simple order. But the groups recognized by chemists such as Mendeleev have an explanation in the substructure of atoms.

**PhET Explorations: Build an Atom**

Build an atom out of protons, neutrons, and electrons, and see how the element, charge, and mass change. Then play a game to test your ideas!

**PhET Interactive Simulation**

Figure 30.63 Build an Atom (http://cnx.org/content/m42618/1.5/build-an-atom_en.jar)

**Glossary**

**angular momentum quantum number:** a quantum number associated with the angular momentum of electrons

**atom:** basic unit of matter, which consists of a central, positively charged nucleus surrounded by negatively charged electrons

**atomic de-excitation:** process by which an atom transfers from an excited electronic state back to the ground state electronic configuration; often occurs by emission of a photon

**atomic excitation:** a state in which an atom or ion acquires the necessary energy to promote one or more of its electrons to electronic states higher in energy than their ground state
**atomic number**: the number of protons in the nucleus of an atom

**Bohr radius**: the mean radius of the orbit of an electron around the nucleus of a hydrogen atom in its ground state

**Brownian motion**: the continuous random movement of particles of matter suspended in a liquid or gas

**cathode-ray tube**: a vacuum tube containing a source of electrons and a screen to view images

**double-slit interference**: an experiment in which waves or particles from a single source impinge upon two slits so that the resulting interference pattern may be observed

**energies of hydrogen-like atoms**: Bohr formula for energies of electron states in hydrogen-like atoms: \[ E_n = -\frac{Z^2}{n^2}E_0(n = 1, 2, 3, \ldots) \]

**energy-level diagram**: a diagram used to analyze the energy level of electrons in the orbits of an atom

**fine structure**: the splitting of spectral lines of the hydrogen spectrum when the spectral lines are examined at very high resolution

**fluorescence**: any process in which an atom or molecule, excited by a photon of a given energy, de-excites by emission of a lower-energy photon

**hologram**: means *entire picture* (from the Greek word *holo*, as in holistic), because the image produced is three dimensional

**holography**: the process of producing holograms

**hydrogen spectrum wavelengths**: the wavelengths of visible light from hydrogen; can be calculated by \[ \lambda = R \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right) \]

**hydrogen-like atom**: any atom with only a single electron

**intrinsic magnetic field**: the magnetic field generated due to the intrinsic spin of electrons

**intrinsic spin**: the internal or intrinsic angular momentum of electrons

**laser**: acronym for light amplification by stimulated emission of radiation

**magnitude of the intrinsic (internal) spin angular momentum**: given by \[ S = \sqrt{s(s + 1)} \frac{\hbar}{2\pi} \]

**metastable**: a state whose lifetime is an order of magnitude longer than the most short-lived states

**orbital angular momentum**: an angular momentum that corresponds to the quantum analog of classical angular momentum

**orbital magnetic field**: the magnetic field generated due to the orbital motion of electrons

**Pauli exclusion principle**: a principle that states that no two electrons can have the same set of quantum numbers; that is, no two electrons can be in the same state

**phosphorescence**: the de-excitation of a metastable state

**planetary model of the atom**: the most familiar model or illustration of the structure of the atom

**population inversion**: the condition in which the majority of atoms in a sample are in a metastable state

**quantum numbers**: the values of quantized entities, such as energy and angular momentum

**Rydberg constant**: a physical constant related to the atomic spectra with an established value of \( 1.097 \times 10^7 \text{ m}^{-1} \)

**shell**: a probability cloud for electrons that has a single principal quantum number

**space quantization**: the fact that the orbital angular momentum can have only certain directions

**spin projection quantum number**: quantum number that can be used to calculate the intrinsic electron angular momentum along the \( z \)-axis

**spin quantum number**: the quantum number that parameterizes the intrinsic angular momentum (or spin angular momentum, or simply spin) of a given particle

**stimulated emission**: emission by atom or molecule in which an excited state is stimulated to decay, most readily caused by a photon of the same energy that is necessary to excite the state

**subshell**: the probability cloud for electrons that has a single angular momentum quantum number \( l \)

**x rays**: a form of electromagnetic radiation

**x-ray diffraction**: a technique that provides the detailed information about crystallographic structure of natural and manufactured materials
Zeeman effect: the effect of external magnetic fields on spectral lines

z-component of spin angular momentum: component of intrinsic electron spin along the z-axis

z-component of the angular momentum: component of orbital angular momentum of electron along the z-axis

Section Summary

30.1 Discovery of the Atom

• Atoms are the smallest unit of elements; atoms combine to form molecules, the smallest unit of compounds.
• The first direct observation of atoms was in Brownian motion.
• Analysis of Brownian motion gave accurate sizes for atoms (10\(^{-10}\) m on average) and a precise value for Avogadro’s number.

30.2 Discovery of the Parts of the Atom: Electrons and Nuclei

• Atoms are composed of negatively charged electrons, first proved to exist in cathode-ray-tube experiments, and a positively charged nucleus.
• All electrons are identical and have a charge-to-mass ratio of \(\frac{q_e}{m_e} = -1.76 \times 10^{11}\) C/kg.
• The positive charge in the nuclei is carried by particles called protons, which have a charge-to-mass ratio of \(\frac{q_p}{m_p} = 9.57 \times 10^7\) C/kg.
• Mass of electron, \(m_e = 9.11 \times 10^{-31}\) kg.
• Mass of proton, \(m_p = 1.67 \times 10^{-27}\) kg.
• The planetary model of the atom pictures electrons orbiting the nucleus in the same way that planets orbit the sun.

30.3 Bohr’s Theory of the Hydrogen Atom

• The planetary model of the atom pictures electrons orbiting the nucleus in the same way that planets orbit the sun. Bohr used the planetary model to develop the first reasonable theory of hydrogen, the simplest atom. Atomic and molecular spectra are quantized, with hydrogen spectrum wavelengths given by the formula

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

where \(\lambda\) is the wavelength of the emitted EM radiation and \(R\) is the Rydberg constant, which has the value \(R = 1.097 \times 10^7\) m\(^{-1}\).

• The constants \(n_i\) and \(n_f\) are positive integers, and \(n_i\) must be greater than \(n_f\).
• Bohr correctly proposed that the energy and radii of the orbits of electrons in atoms are quantized, with energy for transitions between orbits given by

\[
\Delta E = hf = E_i - E_f,
\]

where \(\Delta E\) is the change in energy between the initial and final orbits and \(hf\) is the energy of an absorbed or emitted photon. It is useful to plot orbital energies on a vertical graph called an energy-level diagram.
• Bohr proposed that the allowed orbits are circular and must have quantized orbital angular momentum given by

\[
L = m_e v r_n = n \frac{\hbar}{2\pi}(n = 1, 2, 3 ...),
\]

where \(L\) is the angular momentum, \(r_n\) is the radius of the \(n\)th orbit, and \(\hbar\) is Planck’s constant. For all one-electron (hydrogen-like) atoms, the radius of an orbit is given by

\[
r_n = \frac{n^2 a_B}{Z}(\text{allowed orbits } n = 1, 2, 3, ...),
\]

where \(Z\) is the atomic number of an element (the number of electrons it has when neutral) and \(a_B\) is defined to be the Bohr radius, which is

\[
a_B = \frac{\hbar^2}{4\pi^2 m_e q_e^2} = 0.529 \times 10^{-10}\text{ m}.
\]

• Furthermore, the energies of hydrogen-like atoms are given by

\[
E_n = -\frac{Z^2}{n^2} E_0(n = 1, 2, 3 ...),
\]

where \(E_0\) is the ground-state energy and is given by

\[
E_0 = \frac{2\pi^2 q_e^4 m_e k^2}{\hbar^2} = 13.6 \text{ eV}.
\]

Thus, for hydrogen,
\[ E_n = -\frac{13.6 \text{ eV}}{n^2} (n = 1, 2, 3 ...). \]

- The Bohr Theory gives accurate values for the energy levels in hydrogen-like atoms, but it has been improved upon in several respects.

30.4 X Rays: Atomic Origins and Applications
- X rays are relatively high-frequency EM radiation. They are produced by transitions between inner-shell electron levels, which produce X rays characteristic of the atomic element, or by accelerating electrons.
- X rays have many uses, including medical diagnostics and X-ray diffraction.

30.5 Applications of Atomic Excitations and De-Excitations
- An important atomic process is fluorescence, defined to be any process in which an atom or molecule is excited by absorbing a photon of a given energy and de-excited by emitting a photon of a lower energy.
- Some states live much longer than others and are termed metastable.
- Phosphorescence is the de-excitation of a metastable state.
- Lasers produce coherent single-wavelength EM radiation by stimulated emission, in which a metastable state is stimulated to decay.
- Lasing requires a population inversion, in which a majority of the atoms or molecules are in their metastable state.

30.6 The Wave Nature of Matter Causes Quantization
- Quantization of orbital energy is caused by the wave nature of matter. Allowed orbits in atoms occur for constructive interference of electrons in the orbit, requiring an integral number of wavelengths to fit in an orbit's circumference; that is,
  \[ n\lambda_n = 2\pi r_n (n = 1, 2, 3 ...), \]
  where \( \lambda_n \) is the electron's de Broglie wavelength.
- Owing to the wave nature of electrons and the Heisenberg uncertainty principle, there are no well-defined orbits; rather, there are clouds of probability.
- Bohr correctly proposed that the energy and radii of the orbits of electrons in atoms are quantized, with energy for transitions between orbits given by
  \[ \Delta E = hf = E_i - E_f, \]
  where \( \Delta E \) is the change in energy between the initial and final orbits and \( hf \) is the energy of an absorbed or emitted photon.
- It is useful to plot orbit energies on a vertical graph called an energy-level diagram.
- The allowed orbits are circular, Bohr proposed, and must have quantized orbital angular momentum given by
  \[ L = m_l vr_n = n\frac{h}{2\pi} (n = 1, 2, 3 ...), \]
  where \( L \) is the angular momentum, \( r_n \) is the radius of orbit \( n \), and \( h \) is Planck's constant.

30.7 Patterns in Spectra Reveal More Quantization
- The Zeeman effect—the splitting of lines when a magnetic field is applied—is caused by other quantized entities in atoms.
- Both the magnitude and direction of orbital angular momentum are quantized.
- The same is true for the magnitude and direction of the intrinsic spin of electrons.

30.8 Quantum Numbers and Rules
- Quantum numbers are used to express the allowed values of quantized entities. The principal quantum number \( n \) labels the basic states of a system and is given by
  \[ n = 1, 2, 3, ... \]
- The magnitude of angular momentum is given by
  \[ L = \sqrt{l(l+1)} \frac{h}{2\pi} \quad (l = 0, 1, 2, ..., n-1), \]
  where \( l \) is the angular momentum quantum number. The direction of angular momentum is quantized, in that its component along an axis defined by a magnetic field, called the \( z \)-axis is given by
  \[ L_z = m_l \frac{h}{2\pi} \quad (m_l = -l, -l+1, ..., -1, 0, 1, ..., l-1, l), \]
  where \( L_z \) is the \( z \)-component of the angular momentum and \( m_l \) is the angular momentum projection quantum number. Similarly, the electron’s intrinsic spin angular momentum \( S \) is given by
  \[ S = \sqrt{s(s+1)} \frac{h}{2\pi} \quad (s = 1/2 \text{ for electrons}), \]
  where \( s \) is defined to be the spin quantum number. Finally, the direction of the electron’s spin along the \( z \)-axis is given by
  \[ S_z = m_s \frac{h}{2\pi} \quad (m_s = -\frac{1}{2}, +\frac{1}{2}), \]
  where \( S_z \) is the \( z \)-component of spin angular momentum and \( m_s \) is the spin projection quantum number. Spin projection \( m_s = +1/2 \) is referred to as spin up, whereas \( m_s = -1/2 \) is called spin down. Table 30.1 summarizes the atomic quantum numbers and their allowed values.

30.9 The Pauli Exclusion Principle
- The state of a system is completely described by a complete set of quantum numbers. This set is written as \( \{n, l, m_l, m_s\} \).
The Pauli exclusion principle says that no two electrons can have the same set of quantum numbers; that is, no two electrons can be in the same state.

This exclusion limits the number of electrons in atomic shells and subshells. Each value of \( n \) corresponds to a shell, and each value of \( l \) corresponds to a subshell.

The maximum number of electrons that can be in a subshell is \( 2(2l + 1) \).

The maximum number of electrons that can be in a shell is \( 2n^2 \).

### Conceptual Questions

#### 30.1 Discovery of the Atom

1. Name three different types of evidence for the existence of atoms.

2. Explain why patterns observed in the periodic table of the elements are evidence for the existence of atoms, and why Brownian motion is a more direct type of evidence for their existence.

3. If atoms exist, why can't we see them with visible light?

#### 30.2 Discovery of the Parts of the Atom: Electrons and Nuclei

4. What two pieces of evidence allowed the first calculation of \( m_e \), the mass of the electron?
   
   (a) The ratios \( q_e/m_e \) and \( q_p/m_p \).

   (b) The values of \( q_e \) and \( E_B \).

   (c) The ratio \( q_e/m_e \) and \( q_e \).

   Justify your response.

5. How do the allowed orbits for electrons in atoms differ from the allowed orbits for planets around the sun? Explain how the correspondence principle applies here.

#### 30.3 Bohr’s Theory of the Hydrogen Atom

6. How do the allowed orbits for electrons in atoms differ from the allowed orbits for planets around the sun? Explain how the correspondence principle applies here.

7. Explain how Bohr’s rule for the quantization of electron orbital angular momentum differs from the actual rule.

8. What is a hydrogen-like atom, and how are the energies and radii of its electron orbits related to those in hydrogen?

#### 30.4 X Rays: Atomic Origins and Applications

9. Explain why characteristic x rays are the most energetic in the EM emission spectrum of a given element.

10. Why does the energy of characteristic x rays become increasingly greater for heavier atoms?

11. Observers at a safe distance from an atmospheric test of a nuclear bomb feel its heat but receive none of its copious x rays. Why is air opaque to x rays but transparent to infrared?

12. Lasers are used to burn and read CDs. Explain why a laser that emits blue light would be capable of burning and reading more information than one that emits infrared.

13. Crystal lattices can be examined with x rays but not UV. Why?

14. CT scanners do not detect details smaller than about 0.5 mm. Is this limitation due to the wavelength of x rays? Explain.

#### 30.5 Applications of Atomic Excitations and De-Excitations

15. How do the allowed orbits for electrons in atoms differ from the allowed orbits for planets around the sun? Explain how the correspondence principle applies here.

16. Atomic and molecular spectra are discrete. What does discrete mean, and how are discrete spectra related to the quantization of energy and electron orbits in atoms and molecules?

17. Hydrogen gas can only absorb EM radiation that has an energy corresponding to a transition in the atom, just as it can only emit these discrete energies. When a spectrum is taken of the solar corona, in which a broad range of EM wavelengths are passed through very hot hydrogen gas, the absorption spectrum shows all the features of the emission spectrum. But when such EM radiation passes through room-temperature hydrogen gas, only the Lyman series is absorbed. Explain the difference.

18. Lasers are used to burn and read CDs. Explain why a laser that emits blue light would be capable of burning and reading more information than one that emits infrared.

19. The coating on the inside of fluorescent light tubes absorbs ultraviolet light and subsequently emits visible light. An inventor claims that he is able to do the reverse process. Is the inventor’s claim possible?

20. What is the difference between fluorescence and phosphorescence?

21. How can you tell that a hologram is a true three-dimensional image and that those in 3-D movies are not?

#### 30.6 The Wave Nature of Matter Causes Quantization

22. How is the de Broglie wavelength of electrons related to the quantization of their orbits in atoms and molecules?
30.7 Patterns in Spectra Reveal More Quantization
23. What is the Zeeman effect, and what type of quantization was discovered because of this effect?

30.8 Quantum Numbers and Rules
24. Define the quantum numbers $n$, $l$, $m_l$, $s$, and $m_s$.
25. For a given value of $n$, what are the allowed values of $l$?
26. For a given value of $l$, what are the allowed values of $m_l$? What are the allowed values of $m_l$ for a given value of $n$? Give an example in each case.
27. List all the possible values of $s$ and $m_s$ for an electron. Are there particles for which these values are different? The same?

30.9 The Pauli Exclusion Principle
28. Identify the shell, subshell, and number of electrons for the following: (a) $2p^3$. (b) $4d^9$. (c) $3s^1$. (d) $5g^{16}$.
29. Which of the following are not allowed? State which rule is violated for any that are not allowed. (a) $1p^3$ (b) $2p^8$ (c) $3g^{11}$ (d) $4f^2$
30.1 Discovery of the Atom

1. Using the given charge-to-mass ratios for electrons and protons, and knowing the magnitudes of their charges are equal, what is the ratio of the proton's mass to the electron's? (Note that since the charge-to-mass ratios are given to only three-digit accuracy, your answer may differ from the accepted ratio in the fourth digit.)

2. (a) Calculate the mass of a proton using the charge-to-mass ratio given for it in this chapter and its known charge. (b) How does your result compare with the proton mass given in this chapter?

3. If someone wanted to build a scale model of the atom with a nucleus 1.00 m in diameter, how far away would the nearest electron need to be?

30.2 Discovery of the Parts of the Atom: Electrons and Nuclei

4. Rutherford found the size of the nucleus to be about \(10^{-15}\) m. This implied a huge density. What would this density be for gold?

5. In Millikan's oil-drop experiment, one looks at a small oil drop held motionless between two plates. Take the voltage between the plates to be 2033 V, and the plate separation to be 2.00 cm. The oil drop (of density 0.81 g/cm\(^3\)) has a diameter of 4.0×10\(^{-6}\) m. Find the charge on the drop, in terms of electron units.

6. (a) An aspiring physicist wants to build a scale model of a hydrogen atom for her science fair project. If the atom is 1.00 m in diameter, how big should she try to make the nucleus? (b) How easy will this be to do?

30.3 Bohr's Theory of the Hydrogen Atom

7. By calculating its wavelength, show that the first line in the Lyman series is UV radiation.

8. Find the wavelength of the third line in the Lyman series, and identify the type of EM radiation.

9. Look up the values of the quantities in \(a_B = \frac{\hbar^2}{4\pi^2 m_e kq_e^2}\), and verify that the Bohr radius \(a_B\) is 0.529×10\(^{-10}\) m.

10. Verify that the ground state energy \(E_0\) is 13.6 eV by using \(E_0 = \frac{2\pi^2 q_e^2 m_e k^2}{\hbar^2}\).

11. If a hydrogen atom has its electron in the \(n = 4\) state, how much energy in eV is needed to ionize it?

12. A hydrogen atom in an excited state can be ionized with less energy than when it is in its ground state. What is \(n\) for a hydrogen atom if 0.850 eV of energy can ionize it?

13. Find the radius of a hydrogen atom in the \(n = 2\) state according to Bohr's theory.

14. Show that \((13.6 \text{ eV})/\hbar c = 1.097\times10^7\) m = \(R\) (Rydberg's constant), as discussed in the text.

15. What is the smallest-wavelength line in the Balmer series? Is it in the visible part of the spectrum?

16. Show that the entire Paschen series is in the infrared part of the spectrum. To do this, you only need to consider the shortest wavelength in the series.

17. Do the Balmer and Lyman series overlap? To answer this, calculate the shortest-wavelength Balmer line and the longest-wavelength Lyman line.

18. (a) Which line in the Balmer series is the first one in the UV part of the spectrum? (b) How many Balmer series lines are in the visible part of the spectrum? (c) How many are in the UV?

19. A wavelength of 4.653 µm is observed in a hydrogen spectrum for a transition that ends in the \(n_f = 5\) level. What was \(n_i\) for the initial level of the electron?

20. A singly ionized helium ion has only one electron and is denoted He\(^{+}\). What is the ion's radius in the ground state compared to the Bohr radius of hydrogen atom?

21. A beryllium ion with a single electron (denoted Be\(^{3+}\)) is in an excited state with radius the same as that of the ground state of hydrogen. (a) What is \(n\) for the Be\(^{3+}\) ion? (b) How much energy in eV is needed to ionize the ion from this excited state?

22. Atoms can be ionized by thermal collisions, such as at the high temperatures found in the solar corona. One such ion is C\(^{+5}\), a carbon atom with only a single electron.

(a) By what factor are the energies of its hydrogen-like levels greater than those of hydrogen? (b) What is the wavelength of the first line in this ion's Paschen series? (c) What type of EM radiation is this?

23. Verify Equations \(r_n = \frac{a_B^2}{n^2}\) and \(a_B = \frac{h^2}{4\pi^2 m_e kq_e^2} = 0.529\times10^{-10}\) m using the approach stated in the text. That is, equate the Coulomb and centripetal forces and then insert an expression for velocity from the condition for angular momentum quantization.

24. The wavelength of the four Balmer series lines for hydrogen are found to be 410.3, 434.2, 486.3, and 656.5 nm. What average percentage difference is found between these wavelength numbers and those predicted by \(\frac{1}{\lambda} = \frac{R}{n_i^2} - \frac{1}{n_f^2}\)? It is amazing how well a simple formula (disconnected originally from theory) could duplicate this phenomenon.

30.4 X Rays: Atomic Origins and Applications

25. (a) What is the shortest-wavelength x-ray radiation that can be generated in an x-ray tube with an applied voltage of 50.0 kV? (b) Calculate the photon energy in eV. (c) Explain the relationship of the photon energy to the applied voltage.

26. A color television tube also generates some x rays when its electron beam strikes the screen. What is the shortest wavelength of these x rays, if a 30.0-kV potential is used to accelerate the electrons? (Note that TVs have shielding to prevent these x rays from exposing viewers.)

27. An x-ray tube has an applied voltage of 100 kV. (a) What is the most energetic x-ray photon it can produce? Express your answer in electron volts and joules. (b) Find the wavelength of such an x-ray.

28. The maximum characteristic x-ray photon energy comes from the capture of a free electron into a \(K\) shell vacancy. What is this photon energy in keV for tungsten, assuming the free electron has no initial kinetic energy?

29. What are the approximate energies of the \(K_{\alpha}\) and \(K_{\beta}\) x rays for copper?
30.5 Applications of Atomic Excitations and De-excitations

30. Figure 30.39 shows the energy-level diagram for neon. (a) Verify that the energy of the photon emitted when neon goes from its metastable state to the one immediately below is equal to 1.96 eV. (b) Show that the wavelength of this radiation is 633 nm. (c) What wavelength is emitted when the neon makes a direct transition to its ground state?

31. A helium-neon laser is pumped by electric discharge. What wavelength electromagnetic radiation would be needed to pump it? See Figure 30.39 for energy-level information.

32. Ruby lasers have chromium atoms doped in an aluminum oxide crystal. The energy level diagram for chromium in a ruby is shown in Figure 30.64. What wavelength is emitted by a ruby laser?

<table>
<thead>
<tr>
<th>Shell</th>
<th>Energy Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>0.00 eV</td>
</tr>
<tr>
<td>Metastable</td>
<td>1.79 eV</td>
</tr>
<tr>
<td>Second</td>
<td>2.31 eV</td>
</tr>
<tr>
<td>Third</td>
<td>3.00 eV</td>
</tr>
</tbody>
</table>

Figure 30.64 Chromium atoms in an aluminum oxide crystal have these energy levels, one of which is metastable. This is the basis of a ruby laser. Visible light can pump the atom into an excited state above the metastable state to achieve a population inversion.

33. (a) What energy photons can pump chromium atoms in a ruby laser from the ground state to its second and third excited states? (b) What are the wavelengths of these photons? Verify that they are in the visible part of the spectrum.

34. Some of the most powerful lasers are based on the energy levels of neodymium in solids, such as glass, as shown in Figure 30.65. (a) What average wavelength light can pump the neodymium into the levels above its metastable state? (b) Verify that the 1.17 eV transition produces 1.06 μm radiation.

37. What are the possible values of $m_l$ for an electron in the $n = 4$ state?

38. What, if any, constraints does a value of $m_l = 1$ place on the other quantum numbers for an electron in an atom?

39. (a) Calculate the magnitude of the angular momentum for an $l = 1$ electron. (b) Compare your answer to the value Bohr proposed for the $n = 1$ state.

40. (a) What is the magnitude of the angular momentum for an $l = 1$ electron? (b) Calculate the magnitude of the electron’s spin angular momentum. (c) What is the ratio of these angular momenta?

41. Repeat Exercise 30.40 for $l = 3$.

42. (a) How many angles can $L$ make with the $z$-axis for an $l = 2$ electron? (b) Calculate the value of the smallest angle.

30.9 The Pauli Exclusion Principle

44. (a) How many electrons can be in the $n = 4$ shell? (b) What are its subshells, and how many electrons can be in each?

45. (a) What is the minimum value of $l$ for a subshell that has 11 electrons in it? (b) If this subshell is in the $n = 5$ shell, what is the spectroscopic notation for this atom?

46. (a) If one subshell of an atom has 9 electrons in it, what is the minimum value of $l$? (b) What is the spectroscopic notation for this atom, if this subshell is part of the $n = 3$ shell?

47. (a) List all possible sets of quantum numbers $(n, l, m_l, m_s)$ for the $n = 3$ shell, and determine the number of electrons that can be in the shell and each of its subshells.

(b) Show that the number of electrons in the shell equals $2n^2$ and that the number in each subshell is $2l + 1$.

48. Which of the following spectroscopic notations are not allowed? (a) $5s^1$ (b) $1d^1$ (c) $4s^3$ (d) $3p^7$ (e) $5g^{15}$. State which rule is violated for each that is not allowed.

49. Which of the following spectroscopic notations are allowed (that is, which violate none of the rules regarding values of quantum numbers)? (a) $1s^1$ (b) $1d^1$ (c) $4s^2$ (d) $3p^7$ (e) $6h^{20}$

50. (a) Using the Pauli exclusion principle and the rules relating the allowed values of the quantum numbers $(n, l, m_l, m_s)$, prove that the maximum number of electrons in a subshell is $2n^2$.

(b) In a similar manner, prove that the maximum number of electrons in a shell is $2n^2$.

51. Integrated Concepts

Estimate the density of a nucleus by calculating the density of a proton, taking it to be a sphere 1.2 fm in diameter. Compare your result with the value estimated in this chapter.

52. Integrated Concepts

The electric and magnetic forces on an electron in the CRT in Figure 30.7 are supposed to be in opposite directions. Verify this by determining the direction of each force for the situation shown. Explain how you obtain the directions (that is, identify the rules used).

53. (a) What is the distance between the slits of a diffraction grating that produces a first-order maximum for the first Balmer line at an angle of 20.0°?
(b) At what angle will the fourth line of the Balmer series appear in first order?
(c) At what angle will the second-order maximum be for the first line?

54. Integrated Concepts
A galaxy moving away from the earth has a speed of 0.0100c. What wavelength do we observe for an \( n_i = 7 \) to \( n_f = 2 \) transition for hydrogen in that galaxy?

55. Integrated Concepts
Calculate the velocity of a star moving relative to the earth if you observe a wavelength of 91.0 nm for ionized hydrogen capturing an electron directly into the lowest orbital (that is, \( n_i = \infty \) to \( n_f = 1 \), or a Lyman series transition).

56. Integrated Concepts
In a Millikan oil-drop experiment using a setup like that in Figure 30.9, a 500-V potential difference is applied to plates separated by 2.50 cm. (a) What is the mass of an oil drop having two extra electrons that is suspended motionless by the field between the plates? (b) What is the diameter of the drop, assuming it is a sphere with the density of olive oil?

57. Integrated Concepts
What double-slit separation would produce a first-order maximum at 3.00º for 25.0-keV x rays? The small answer indicates that the wave character of x rays is best determined by having them interact with very small objects such as atoms and molecules.

58. Integrated Concepts
In a laboratory experiment designed to duplicate Thomson’s determination of \( q_e/m_e \), a beam of electrons having a velocity of 6.00×10^7 m/s enters a 5.00×10^{-3} T magnetic field. The beam moves perpendicular to the field in a path having a 6.80-cm radius of curvature. Determine \( q_e/m_e \) from these observations, and compare the result with the known value.

59. Integrated Concepts
Find the value of \( l \), the orbital angular momentum quantum number, for the moon around the earth. The extremely large value obtained implies that it is impossible to tell the difference between adjacent quantized orbits for macroscopic objects.

60. Integrated Concepts
Particles called muons exist in cosmic rays and can be created in particle accelerators. Muons are very similar to electrons, having the same charge and spin, but they have a mass 207 times greater. When muons are captured by an atom, they orbit just like an electron but with a smaller radius, since the mass in \( d_B = \frac{\hbar^2}{4\pi^2 m_e k q_e} = 0.529\times10^{-10} \) m is 207 \( m_e \).

(a) Calculate the radius of the \( n = 1 \) orbit for a muon in a uranium ion (\( Z = 92 \)).
(b) Compare this with the 7.5-fm radius of a uranium nucleus. Note that since the muon orbits inside the electron, it falls into a hydrogen-like orbit.

61. Integrated Concepts
Calculate the minimum amount of energy in joules needed to create a population inversion in a helium-neon laser containing 1.00×10^{-4} moles of neon.

62. Integrated Concepts
A carbon dioxide laser used in surgery emits infrared radiation with a wavelength of 10.6 \( \mu \)m. In 1.00 ms, this laser raised the temperature of 1.00 cm^3 of flesh to 100ºC and evaporated it.

(a) How many photons were required? You may assume flesh has the same heat of vaporization as water. (b) What was the minimum power output during the flash?

63. Integrated Concepts
Suppose an MRI scanner uses 100-MHz radio waves.

(a) Calculate the photon energy.
(b) How does this compare to typical molecular binding energies?

64. Integrated Concepts
(a) An excimer laser used for vision correction emits 193-nm UV. Calculate the photon energy in eV.
(b) These photons are used to evaporate corneal tissue, which is very similar to water in its properties. Calculate the amount of energy needed per molecule of water to make the phase change from liquid to gas. That is, divide the heat of vaporization in kJ/kg by the number of water molecules in a kilogram.

(c) Convert this to eV and compare to the photon energy. Discuss the implications.

65. Integrated Concepts
A neighboring galaxy rotates on its axis so that stars on one side move toward us as fast as 200 km/s, while those on the other side move away as fast as 200 km/s. This causes the EM radiation we receive to be Doppler shifted by velocities over the entire range of ±200 km/s. What range of wavelengths will we observe for the 656.0-nm line in the Balmer series of hydrogen emitted by stars in this galaxy. (This is called line broadening.)

66. Integrated Concepts
A pulsar is a rapidly spinning remnant of a supernova. It rotates on its axis, sweeping hydrogen along with it so that hydrogen on one side moves toward us as fast as 50.0 km/s, while that on the other side moves away as fast as 50.0 km/s. This means that the EM radiation we receive will be Doppler shifted over a range of ±50.0 km/s. What range of wavelengths will we observe for the 91.20-nm line in the Lyman series of hydrogen? (Such line broadening is observed and actually provides part of the evidence for rapid rotation.)

67. Integrated Concepts
Prove that the velocity of charged particles moving along a straight path through perpendicular electric and magnetic fields is \( v = E/B \). Thus crossed electric and magnetic fields can be used as a velocity selector independent of the charge and mass of the particle involved.

68. Unreasonable Results
(a) What voltage must be applied to an X-ray tube to obtain 0.0100-fm wavelength X-rays for use in exploring the details of nuclei? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

69. Unreasonable Results
A student in a physics laboratory observes a hydrogen spectrum with a diffraction grating for the purpose of measuring the wavelengths of the emitted radiation. In the spectrum, she observes a yellow line and finds its wavelength to be 589 nm. (a) Assuming this is part of the Balmer series, determine \( n_i \), the principal quantum number of the initial state.

(b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

70. Construct Your Own Problem
The solar corona is so hot that most atoms in it are ionized. Consider a hydrogen-like atom in the corona that has only a single electron. Construct a problem in which you calculate selected spectral energies and wavelengths of the Lyman, Balmer, or other series of this atom that could be used to identify its presence in a very hot gas. You will need to
choose the atomic number of the atom, identify the element, and choose which spectral lines to consider.

71. Construct Your Own Problem

Consider the Doppler-shifted hydrogen spectrum received from a rapidly receding galaxy. Construct a problem in which you calculate the energies of selected spectral lines in the Balmer series and examine whether they can be described with a formula like that in the equation

$$\frac{1}{\lambda} = R \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right),$$

but with a different constant $R$. 

$\lambda$
31 RADIOACTIVITY AND NUCLEAR PHYSICS

Learning Objectives

31.1. Nuclear Radioactivity
- Explain nuclear radiation.
- Explain the types of radiation—alpha emission, beta emission, and gamma emission.
- Explain the ionization of radiation in an atom.
- Define the range of radiation.

31.2. Radiation Detection and Detectors
- Explain the working principle of a Geiger tube.
- Define and discuss radiation detectors.

31.3. Substructure of the Nucleus
- Define and discuss the nucleus in an atom.
- Define atomic number.
- Define and discuss isotopes.
- Calculate the density of the nucleus.
- Explain nuclear force.

31.4. Nuclear Decay and Conservation Laws
- Define and discuss nuclear decay.
- State the conservation laws.
- Explain parent and daughter nucleus.
- Calculate the energy emitted during nuclear decay.

31.5. Half-Life and Activity
- Define half-life.
- Define dating.
- Calculate age of old objects by radioactive dating.

31.6. Binding Energy
- Define and discuss binding energy.
- Calculate the binding energy per nucleon of a particle.

31.7. Tunneling
- Define and discuss tunneling.
- Define potential barrier.
- Explain quantum tunneling.
**Introduction to Radioactivity and Nuclear Physics**

There is an ongoing quest to find substructures of matter. At one time, it was thought that atoms would be the ultimate substructure, but just when the first direct evidence of atoms was obtained, it became clear that they have a substructure and a tiny nucleus. The nucleus itself has spectacular characteristics. For example, certain nuclei are unstable, and their decay emits radiations with energies millions of times greater than atomic energies. Some of the mysteries of nature, such as why the core of the earth remains molten and how the sun produces its energy, are explained by nuclear phenomena. The exploration of radioactivity and the nucleus revealed fundamental and previously unknown particles, forces, and conservation laws. That exploration has evolved into a search for further underlying structures, such as quarks. In this chapter, the fundamentals of nuclear radioactivity and the nucleus are explored. The following two chapters explore the more important applications of nuclear physics in the field of medicine. We will also explore the basics of what we know about quarks and other substructures smaller than nuclei.

### 31.1 Nuclear Radioactivity

The discovery and study of nuclear radioactivity quickly revealed evidence of revolutionary new physics. In addition, uses for nuclear radiation also emerged quickly—for example, people such as Ernest Rutherford used it to determine the size of the nucleus and devices were painted with radon-doped paint to make them glow in the dark (see Figure 31.2). We therefore begin our study of nuclear physics with the discovery and basic features of nuclear radioactivity.

**Figure 31.2** The dials of this World War II aircraft glow in the dark, because they are painted with radium-doped phosphorescent paint. It is a poignant reminder of the dual nature of radiation. Although radium paint dials are conveniently visible day and night, they emit radon, a radioactive gas that is hazardous and is not directly sensed. (credit: U.S. Air Force Photo)

**Discovery of Nuclear Radioactivity**

In 1896, the French physicist Antoine Henri Becquerel (1852–1908) accidentally found that a uranium-rich mineral called pitchblende emits invisible, penetrating rays that can darken a photographic plate enclosed in an opaque envelope. The rays therefore carry energy; but amazingly, the pitchblende emits them continuously without any energy input. This is an apparent violation of the law of conservation of energy, one that we now understand is due to the conversion of a small amount of mass into energy, as related in Einstein’s famous equation $E = mc^2$. It was soon evident that Becquerel’s rays originate in the nuclei of the atoms and have other unique characteristics. The emission of these rays is called nuclear radioactivity or simply radioactive. The rays themselves are called nuclear radiation. A nucleus that spontaneously destroys part of its mass to emit radiation is said to decay (a term also used to describe the emission of radiation by atoms in excited states). A substance or object that emits nuclear radiation is said to be radioactive.

Two types of experimental evidence imply that Becquerel’s rays originate deep in the heart (or nucleus) of an atom. First, the radiation is found to be associated with certain elements, such as uranium. Radiation does not vary with chemical state—that is, uranium is radioactive whether it is in the form of an element or compound. In addition, radiation does not vary with temperature, pressure, or ionization state of the uranium atom. Since all of these factors affect electrons in an atom, the radiation cannot come from electron transitions, as atomic spectra do. The huge energy emitted during each event is the second piece of evidence that the radiation cannot be atomic. Nuclear radiation has energies of the order of $10^6$ eV per event, which is much greater than the typical atomic energies (a few eV), such as that observed in spectra and chemical reactions, and more than ten times as high as the most energetic characteristic x rays. Becquerel did not vigorously pursue his discovery for very long. In 1898, Marie Curie (1867–1934), then a graduate student married the already well-known French physicist Pierre Curie (1859–1906), began her doctoral study of Becquerel’s rays. She and her husband soon discovered two new radioactive elements, which she named polonium (after her native land) and radium (because it radiates). These two new elements filled holes in the periodic table and, further, displayed much higher levels of radioactivity per gram of material than uranium. Over a period of four years, working under poor conditions and spending their own funds, the Curies processed more than a ton of uranium ore to isolate a gram of radium salt. Radium became highly sought after, because it was about two million times as radioactive as uranium. Curie’s radium salt glowed visibly from the radiation that took its toll on them and other unaware researchers. Shortly after completing her Ph.D., both Curies and Becquerel shared the 1903 Nobel Prize in physics for their work on radioactivity. Pierre was killed in a horse cart accident in 1906, but Marie continued her study of radioactivity for nearly 30 more years. Awarded the 1911 Nobel Prize in chemistry for her discovery of two new elements, she remains the only person to win Nobel Prizes in physics and chemistry. Marie’s radioactive fingerprints on some pages of her notebooks can still expose film, and she suffered from radiation-induced lesions. She died of leukemia likely caused by radiation, but she was active in research almost until her death in 1934. The following year, her daughter and son-in-law, Irene and Frederic Joliot-Curie, were awarded the Nobel Prize in chemistry for their discovery of artificially induced radiation, adding to a remarkable family legacy.

**Alpha, Beta, and Gamma**

Research begun by people such as New Zealander Ernest Rutherford soon after the discovery of nuclear radiation indicated that different types of rays are emitted. Eventually, three types were distinguished and named alpha ($\alpha$), beta ($\beta$), and gamma ($\gamma$), because, like x-rays, their identities were initially unknown. Figure 31.3 shows what happens if the rays are passed through a magnetic field. The $\gamma$’s are unaffected, while the $\alpha$’s and
\(\beta\) s are deflected in opposite directions, indicating the \(\alpha\) s are positive, the \(\beta\) s negative, and the \(\gamma\) s uncharged. Rutherford used both magnetic and electric fields to show that \(\alpha\) s have a positive charge twice the magnitude of an electron, or \(+2 | q_e |\). In the process, he found the \(\alpha\) s charge to mass ratio to be several thousand times smaller than the electron’s. Later on, Rutherford collected \(\alpha\) s from a radioactive source and passed an electric discharge through them, obtaining the spectrum of recently discovered helium gas. Among many important discoveries made by Rutherford and his collaborators was the proof that \(\alpha\) radiation is the emission of a helium nucleus. Rutherford won the Nobel Prize in chemistry in 1908 for his early work. He continued to make important contributions until his death in 1934.

Figure 31.3 Alpha, beta, and gamma rays are passed through a magnetic field on the way to a phosphorescent screen. The \(\alpha\) s and \(\beta\) s bend in opposite directions, while the \(\gamma\) s are unaffected, indicating a positive charge for \(\alpha\) s, negative for \(\beta\) s, and neutral for \(\gamma\) s. Consistent results are obtained with electric fields. Collection of the radiation offers further confirmation from the direct measurement of excess charge.

Other researchers had already proved that \(\beta\) s are negative and have the same mass and same charge-to-mass ratio as the recently discovered electron. By 1902, it was recognized that \(\beta\) radiation is the emission of an electron. Although \(\beta\) s are electrons, they do not exist in the nucleus before it decays and are not ejected atomic electrons—the electron is created in the nucleus at the instant of decay.

Since \(\gamma\) s remain unaffected by electric and magnetic fields, it is natural to think they might be photons. Evidence for this grew, but it was not until 1914 that this was proved by Rutherford and collaborators. By scattering \(\gamma\) radiation from a crystal and observing interference, they demonstrated that \(\gamma\) radiation is the emission of a high-energy photon by a nucleus. In fact, \(\gamma\) radiation comes from the de-excitation of a nucleus, just as an \(x\) ray comes from the de-excitation of an atom. The names ”\(\gamma\) ray” and ”\(x\) ray” identify the source of the radiation. At the same energy, \(\gamma\) rays and \(x\) rays are otherwise identical.

<table>
<thead>
<tr>
<th>Type of Radiation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)-Particles</td>
<td>A sheet of paper, a few cm of air, fractions of a mm of tissue</td>
</tr>
<tr>
<td>(\beta)-Particles</td>
<td>A thin aluminum plate, or tens of cm of tissue</td>
</tr>
<tr>
<td>(\gamma) Rays</td>
<td>Several cm of lead or meters of concrete</td>
</tr>
</tbody>
</table>

**Ionization and Range**

Two of the most important characteristics of \(\alpha\), \(\beta\), and \(\gamma\) rays were recognized very early. All three types of nuclear radiation produce ionization in materials, but they penetrate different distances in materials—that is, they have different ranges. Let us examine why they have these characteristics and what are some of the consequences.

Like \(x\) rays, nuclear radiation in the form of \(\alpha\) s, \(\beta\) s, and \(\gamma\) s has enough energy per event to ionize atoms and molecules in any material. The energy emitted in various nuclear decays ranges from a few keV to more than 10 MeV, while only a few eV are needed to produce ionization.

The effects of \(x\) rays and nuclear radiation on biological tissues and other materials, such as solid state electronics, are directly related to the ionization they produce. All of them, for example, can damage electronics or kill cancer cells. In addition, methods for detecting \(x\) rays and nuclear radiation are based on ionization, directly or indirectly. All of them can ionize the air between the plates of a capacitor, for example, causing it to discharge. This is the basis of inexpensive personal radiation monitors, such as pictured in Figure 31.4. Apart from \(\alpha\), \(\beta\), and \(\gamma\), there are other
forms of nuclear radiation as well, and these also produce ionization with similar effects. We define ionizing radiation as any form of radiation that produces ionization whether nuclear in origin or not, since the effects and detection of the radiation are related to ionization.

![Figure 31.4](credit: L. Chang, Wikimedia Commons)

The range of radiation is defined to be the distance it can travel through a material. Range is related to several factors, including the energy of the radiation, the material encountered, and the type of radiation (see Figure 31.5). The higher the energy, the greater the range, all other factors being the same. This makes good sense, since radiation loses its energy in materials primarily by producing ionization in them, and each ionization of an atom or a molecule requires energy that is removed from the radiation. The amount of ionization is, thus, directly proportional to the energy of the particle of radiation, as is its range.

![Figure 31.5](a) Greater energy means greater range. (b) Radiation has a smaller range in materials with high electron density. (c) Alphas have the smallest range, betas have a greater range, and gammas penetrate the farthest.

Radiation can be absorbed or shielded by materials, such as the lead aprons dentists drape on us when taking x rays. Lead is a particularly effective shield compared with other materials, such as plastic or air. How does the range of radiation depend on material? Ionizing radiation interacts best with charged particles in a material. Since electrons have small masses, they most readily absorb the energy of the radiation in collisions. The greater the density of a material and, in particular, the greater the density of electrons within a material, the smaller the range of radiation.

Collisions

Conservation of energy and momentum often results in energy transfer to a less massive object in a collision. This was discussed in detail in Work, Energy, and Energy Resources, for example.

Different types of radiation have different ranges when compared at the same energy and in the same material. Alphas have the shortest range, betas penetrate farther, and gammas have the greatest range. This is directly related to charge and speed of the particle or type of radiation. At a given energy, each \( \alpha \), \( \beta \), or \( \gamma \) will produce the same number of ionizations in a material (each ionization requires a certain amount of energy on average). The more readily the particle produces ionization, the more quickly it will lose its energy. The effect of charge is as follows: The \( \alpha \) has a charge of \( +2q_e \), the \( \beta \) has a charge of \( -2q_e \), and the \( \gamma \) is uncharged. The electromagnetic force exerted by the \( \alpha \) is thus twice as strong as that exerted by the \( \beta \) and it is more likely to produce ionization. Although chargeless, the \( \gamma \) does interact weakly because it is an electromagnetic wave, but it is less likely to produce ionization in any encounter. More quantitatively, the change in momentum \( \Delta p \) given to a particle in the material is \( \Delta p = F \Delta t \), where \( F \) is the force the \( \alpha \), \( \beta \), or \( \gamma \) exerts over a time \( \Delta t \). The smaller the charge, the smaller is \( F \) and the smaller is the momentum (and energy) lost. Since the speed of alphas is about 5% to 10% of the speed of light, classical (non-relativistic) formulas apply.

The speed at which they travel is the other major factor affecting the range of \( \alpha \)s, \( \beta \)s, and \( \gamma \)s. The faster they move, the less time they spend in the vicinity of an atom or a molecule, and the less likely they are to interact. Since \( \alpha \) s and \( \beta \) s are particles with mass (helium nuclei and electrons,
respectively), their energy is kinetic, given classically by $\frac{1}{2}mv^2$. The mass of the $\beta$ particle is thousands of times less than that of the $\alpha$ s, so that $\beta$ s must travel much faster than $\alpha$ s to have the same energy. Since $\beta$ s move faster (most at relativistic speeds), they have less time to interact than $\alpha$ s. Gamma rays are photons, which must travel at the speed of light. They are even less likely to interact than a $\beta$, since they spend even less time near a given atom (and they have no charge). The range of $\gamma$ s is thus greater than the range of $\beta$ s.

Alpha radiation from radioactive sources has a range much less than a millimeter of biological tissues, usually not enough to even penetrate the dead layers of our skin. On the other hand, the same $\alpha$ radiation can penetrate a few centimeters of air, so mere distance from a source prevents $\alpha$ radiation from reaching us. This makes $\alpha$ radiation relatively safe for our body compared to $\beta$ and $\gamma$ radiation. Typical $\beta$ radiation can penetrate a few millimeters of tissue or about a meter of air. Beta radiation is thus hazardous even when not ingested. The range of $\beta$ s in lead is about a millimeter, and so it is easy to store $\beta$ sources in lead radiation-proof containers. Gamma rays have a much greater range than either $\alpha$ s or $\beta$ s. In fact, if a given thickness of material, like a lead brick, absorbs 90% of the $\gamma$ s, then a second lead brick will only absorb 90% of what got through the first. Thus, $\gamma$ s do not have a well-defined range; we can only cut down the amount that gets through. Typically, $\gamma$ s can penetrate many meters of air, go right through our bodies, and are effectively shielded (that is, reduced in intensity to acceptable levels) by many centimeters of lead. One benefit of $\gamma$ s is that they can be used as radioactive tracers (see Figure 31.6).

![Figure 31.6](http://cnx.org/content/m42623/1.6/beta-decay_en.jar)

**Figure 31.6** This image of the concentration of a radioactive tracer in a patient’s body reveals where the most active bone cells are, an indication of bone cancer. A short-lived radioactive substance that locates itself selectively is given to the patient, and the radiation is measured with an external detector. The emitted $\gamma$ radiation has a sufficient range to leave the body—the range of $\alpha$ s and $\beta$ s is too small for them to be observed outside the patient. (credit: Kieran Maher, Wikimedia Commons)

**31.2 Radiation Detection and Detectors**

It is well known that ionizing radiation affects us but does not trigger nerve impulses. Newspapers carry stories about unsuspecting victims of radiation poisoning who fall ill with radiation sickness, such as burns and blood count changes, but who never felt the radiation directly. This makes the detection of radiation by instruments more than an important research tool. This section is a brief overview of radiation detection and some of its applications.

**Human Application**

The first direct detection of radiation was Becquerel’s fogged photographic plate. Photographic film is still the most common detector of ionizing radiation, being used routinely in medical and dental x rays. Nuclear radiation is also captured on film, such as seen in Figure 31.8. The mechanism for film exposure by ionizing radiation is similar to that by photons. A quantum of energy interacts with the emulsion and alters it chemically, thus exposing the film. The quantum come from an $\alpha$-particle, $\beta$-particle, or photon, provided it has more than the few eV of energy needed to induce the chemical change (as does all ionizing radiation). The process is not 100% efficient, since not all incident radiation interacts and not all interactions produce the chemical change. The amount of film darkening is related to exposure, but the darkening also depends on the type of radiation, so that absorbers and other devices must be used to obtain energy, charge, and particle-identification information.
Another very common radiation detector is the Geiger tube. The clicking and buzzing sound we hear in dramatizations and documentaries, as well as in our own physics labs, is usually an audio output of events detected by a Geiger counter. These relatively inexpensive radiation detectors are based on the simple and sturdy Geiger tube, shown schematically in Figure 31.9(b). A conducting cylinder with a wire along its axis is filled with an insulating gas so that a voltage applied between the cylinder and wire produces almost no current. Ionizing radiation passing through the tube produces free ion pairs that are attracted to the wire and cylinder, forming a current that is detected as a count. The word count implies that there is no information on energy, charge, or type of radiation with a simple Geiger counter. They do not detect every particle, since some radiation can pass through without producing enough ionization to be detected. However, Geiger counters are very useful in producing a prompt output that reveals the existence and relative intensity of ionizing radiation.

Another radiation detection method records light produced when radiation interacts with materials. The energy of the radiation is sufficient to excite atoms in a material that may fluoresce, such as the phosphor used by Rutherford's group. Materials called scintillators use a more complex collaborative process to convert radiation energy into light. Scintillators may be liquid or solid, and they can be very efficient. Their light output can provide information about the energy, charge, and type of radiation. Scintillator light flashes are very brief in duration, enabling the detection of a huge number of particles in short periods of time. Scintillator detectors are used in a variety of research and diagnostic applications. Among these are the detection by satellite-mounted equipment of the radiation from distant galaxies, the analysis of radiation from a person indicating body burdens, and the detection of exotic particles in accelerator laboratories.
Light from a scintillator is converted into electrical signals by devices such as the photomultiplier tube shown schematically in Figure 31.10. These tubes are based on the photoelectric effect, which is multiplied in stages into a cascade of electrons, hence the name photomultiplier. Light entering the photomultiplier strikes a metal plate, ejecting an electron that is attracted by a positive potential difference to the next plate, giving it enough energy to eject two or more electrons, and so on. The final output current can be made proportional to the energy of the light entering the tube, which is in turn proportional to the energy deposited in the scintillator. Very sophisticated information can be obtained with scintillators, including energy, charge, particle identification, direction of motion, and so on.

Solid-state radiation detectors convert ionization produced in a semiconductor (like those found in computer chips) directly into an electrical signal. Semiconductors can be constructed that do not conduct current in one particular direction. When a voltage is applied in that direction, current flows only when ionization is produced by radiation, similar to what happens in a Geiger tube. Further, the amount of current in a solid-state detector is closely related to the energy deposited and, since the detector is solid, it can have a high efficiency (since ionizing radiation is stopped in a shorter distance in solids fewer particles escape detection). As with scintillators, very sophisticated information can be obtained from solid-state detectors.

31.3 Substructure of the Nucleus

What is inside the nucleus? Why are some nuclei stable while others decay? (See Figure 31.12.) Why are there different types of decay ($\alpha$, $\beta$, and $\gamma$)? Why are nuclear decay energies so large? Pursuing natural questions like these has led to far more fundamental discoveries than you might imagine.
We have already identified protons as the particles that carry positive charge in the nuclei. However, there are actually two types of particles in the nuclei—the proton and the neutron, referred to collectively as nucleons, the constituents of nuclei. As its name implies, the neutron is a neutral particle (\( q = 0 \)) that has nearly the same mass and intrinsic spin as the proton. Table 31.2 compares the masses of protons, neutrons, and electrons. Note how close the proton and neutron masses are, but the neutron is slightly more massive once you look past the third digit. Both nucleons are much more massive than an electron. In fact, \( m_p = 1.836 \times 10^{-31} \) kg (as noted in Medical Applications of Nuclear Physics) and \( m_n = 1.839 \times 10^{-27} \) kg. 

Table 31.2 also gives masses in terms of mass units that are more convenient than kilograms on the atomic and nuclear scale. The first of these is the unified atomic mass unit (u), defined as

\[
1 \text{ u} = 1.6605 \times 10^{-27} \text{ kg}.
\]

This unit is defined so that a neutral carbon \( ^{12}C \) atom has a mass of exactly 12 u. Masses are also expressed in units of MeV/c². These units are very convenient when considering the conversion of mass into energy (and vice versa), as is so prominent in nuclear processes. Using \( E = mc^2 \) and units of \( m \) in MeV/c², we find that \( c^2 \) cancels and \( E \) comes out conveniently in MeV. For example, if the rest mass of a proton is converted entirely into energy, then

\[
E = mc^2 = (938.27 \text{ MeV/c}^2) c^2 = 938.27 \text{ MeV}.
\]

It is useful to note that 1 u of mass converted to energy produces 931.5 MeV, or

\[
1 \text{ u} = 931.5 \text{ MeV/c}^2.
\]

All properties of a nucleus are determined by the number of protons and neutrons it has. A specific combination of protons and neutrons is called a nuclide and is a unique nucleus. The following notation is used to represent a particular nuclide:

\[
\frac{A}{Z}X_N,
\]

where the symbols \( A \), \( X \), \( Z \), and \( N \) are defined as follows: The number of protons in a nucleus is the atomic number \( Z \), as defined in Medical Applications of Nuclear Physics. \( X \) is the symbol for the element, such as Ca for calcium. However, once \( Z \) is known, the element is known; hence, \( Z \) and \( X \) are redundant. For example, \( Z = 20 \) is always calcium, and calcium always has \( Z = 20 \). \( N \) is the number of neutrons in a nucleus. In the notation for a nuclide, the subscript \( N \) is usually omitted. The symbol \( A \) is defined as the number of nucleons or the total number of protons and neutrons,

\[
A = N + Z,
\]

where \( A \) is also called the mass number. This name for \( A \) is logical; the mass of an atom is nearly equal to the mass of its nucleus, since electrons have so little mass. The mass of the nucleus turns out to be nearly equal to the sum of the masses of the protons and neutrons in it, which is proportional to \( A \). In this context, it is particularly convenient to express masses in units of u. Both protons and neutrons have masses close to 1 u, and so the mass of an atom is close to \( A \) u. For example, in an oxygen nucleus with eight protons and eight neutrons, \( A = 16 \), and its mass is 16 u. As noticed, the unified atomic mass unit is defined so that a neutral carbon atom (actually a \( ^{12}C \) atom) has a mass of exactly 12 u. Carbon was chosen as the standard, partly because of its importance in organic chemistry (see Appendix A).

<table>
<thead>
<tr>
<th>Particle</th>
<th>Symbol</th>
<th>kg</th>
<th>u</th>
<th>MeV/c²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proton</td>
<td>( p )</td>
<td>( 1.67262 \times 10^{-27} )</td>
<td>( 1.007276 )</td>
<td>( 938.27 )</td>
</tr>
<tr>
<td>Neutron</td>
<td>( n )</td>
<td>( 1.67493 \times 10^{-27} )</td>
<td>( 1.008665 )</td>
<td>( 939.57 )</td>
</tr>
<tr>
<td>Electron</td>
<td>( e )</td>
<td>( 9.1094 \times 10^{-31} )</td>
<td>( 0.00054858 )</td>
<td>( 0.511 )</td>
</tr>
</tbody>
</table>

Let us look at a few examples of nuclides expressed in the \( \frac{A}{Z}X_N \) notation. The nucleus of the simplest atom, hydrogen, is a single proton, or \( ^{1}_1\text{H} \) (the zero for no neutrons is often omitted). To check this symbol, refer to the periodic table—you see that the atomic number \( Z \) of hydrogen is 1.
Since you are given that there are no neutrons, the mass number $A$ is also 1. Suppose you are told that the helium nucleus or $\alpha$ particle has two protons and two neutrons. You can then see that it is written $^4_2\text{He}_2$. There is a scarce form of hydrogen found in nature called deuterium; its nucleus has one proton and one neutron and, hence, twice the mass of common hydrogen. The symbol for deuterium is, thus, $^2_1\text{H}_1$ (sometimes D is used, as for deuterated water $\text{D}_2\text{O}$). An even rarer—and radioactive—form of hydrogen is called tritium, since it has a single proton and two neutrons, and it is written $^3_1\text{H}_2$. These three varieties of hydrogen have nearly identical chemistries, but the nuclei differ greatly in mass, stability, and other characteristics. Nuclei (such as those of hydrogen) having the same $Z$ and different $N$'s are defined to be isotopes of the same element.

There is some redundancy in the symbols $A$, $X$, $Z$, and $N$. If the element $X$ is known, then $Z$ can be found in a periodic table and is always the same for a given element. If both $A$ and $X$ are known, then $N$ can also be determined (first find $Z$; then, $N = A - Z$). Thus the simpler notation for nuclides is $^A_X$.

A variety of experiments indicate that a nucleus behaves something like a tightly packed ball of nucleons, as illustrated in Figure 31.13. These nucleons have large kinetic energies and, thus, move rapidly in very close contact. Nucleons can be separated by a large force, such as in a collision with another nucleus, but resist strongly being pushed closer together. The most compelling evidence that nucleons are closely packed in a nucleus is that the radius of a nucleus, $r$, is found to be given approximately by

$$r = r_0 A^{1/3},$$

where $r_0 = 1.2 \text{ fm}$ and $A$ is the mass number of the nucleus. Note that $r^3 \propto A$. Since many nuclei are spherical, and the volume of a sphere is $V = (4/3)\pi r^3$, we see that $V \propto A$ —that is, the volume of a nucleus is proportional to the number of nucleons in it. This is what would happen if you pack nucleons so closely that there is no empty space between them.

![Proton](Proton.png) ![Neutron](Neutron.png)

Figure 31.13 A model of the nucleus.

Nucleons are held together by nuclear forces and resist both being pulled apart and pushed inside one another. The volume of the nucleus is the sum of the volumes of the nucleons in it, here shown in different colors to represent protons and neutrons.

**Example 31.1 How Small and Dense Is a Nucleus?**

(a) Find the radius of an iron-56 nucleus. (b) Find its approximate density in kg/m$^3$, approximating the mass of $^{56}\text{Fe}$ to be 56 u.

**Strategy and Concept**

(a) Finding the radius of $^{56}\text{Fe}$ is a straightforward application of $r = r_0 A^{1/3}$, given $A = 56$. (b) To find the approximate density, we assume the nucleus is spherical (this one actually is), calculate its volume using the radius found in part (a), and then find its density from $\rho = m/V$.

Finally, we will need to convert density from units of u/fm$^3$ to kg/m$^3$.

**Solution**

(a) The radius of a nucleus is given by

$$r = r_0 A^{1/3}.$$  

Substituting the values for $r_0$ and $A$ yields

$$r = (1.2 \text{ fm})(56)^{1/3} = (1.2 \text{ fm})(3.83)$$

$$= 4.6 \text{ fm}.$$  

(b) Density is defined to be $\rho = m/V$, which for a sphere of radius $r$ is

$$\rho = \frac{m}{V} = \frac{m}{(4/3)\pi r^3}.$$
Substituting known values gives

\[
\rho = \frac{56 \text{ u}}{(1.33)(3.14)(4.6 \text{ fm})^3} = 0.138 \text{ u/fm}^3.
\] (31.11)

Converting to units of kg/m³, we find

\[
\rho = (0.138 \text{ u/fm}^3)(1.66 \times 10^{-27} \text{ kg/u})\left(\frac{1 \text{ fm}}{10^{-15} \text{ m}}\right) = 2.3 \times 10^{17} \text{ kg/m}^3.
\] (31.12)

Discussion

(a) The radius of this medium-sized nucleus is found to be approximately 4.6 fm, and so its diameter is about 10 fm, or 10⁻¹⁴ m. In our discussion of Rutherford’s discovery of the nucleus, we noticed that it is about 10⁻¹⁵ m in diameter (which is for lighter nuclei), consistent with this result to an order of magnitude. The nucleus is much smaller in diameter than the typical atom, which has a diameter of the order of 10⁻¹⁰ m.

(b) The density found here is so large as to cause disbelief. It is consistent with earlier discussions we have had about the nucleus being very small and containing nearly all of the mass of the atom. Nuclear densities, such as found here, are about 2 × 10¹⁴ times greater than that of water, which has a density of “only” 10³ kg/m³. One cubic meter of nuclear matter, such as found in a neutron star, has the same mass as a cube of water 61 km on a side.

Nuclear Forces and Stability

What forces hold a nucleus together? The nucleus is very small and its protons, being positive, exert tremendous repulsive forces on one another. (The Coulomb force increases as charges get closer, since it is proportional to 1/r², even at the tiny distances found in nuclei.) The answer is that two previously unknown forces hold the nucleus together and make it into a tightly packed ball of nucleons. These forces are called the weak and strong nuclear forces. Nuclear forces are so short ranged that they fall to zero strength when nucleons are separated by only a few fm. However, like glue, they are strongly attracted when the nucleons get close to one another. The strong nuclear force is about 100 times more attractive than the repulsive EM force, easily holding the nucleons together. Nuclear forces become extremely repulsive if the nucleons get too close, making nucleons strongly resist being pushed inside one another, something like ball bearings.

The fact that nuclear forces are very strong is responsible for the very large energies emitted in nuclear decay. During decay, the forces do work, and since work is force times the distance (W = Fd cos θ), a large force can result in a large emitted energy. In fact, we know that there are two distinct nuclear forces because of the different types of nuclear decay—the strong nuclear force is responsible for α decay, while the weak nuclear force is responsible for β decay.

The many stable and unstable nuclei we have explored, and the hundreds we have not discussed, can be arranged in a table called the chart of the nuclides, a simplified version of which is shown in Figure 31.14. Nuclides are located on a plot of N versus Z. Examination of a detailed chart of the nuclides reveals patterns in the characteristics of nuclei, such as stability, abundance, and types of decay, analogous to but more complex than the systematics in the periodic table of the elements.
In principle, a nucleus can have any combination of protons and neutrons, but Figure 31.14 shows a definite pattern for those that are stable. For low-mass nuclei, there is a strong tendency for $N$ and $Z$ to be nearly equal. This means that the nuclear force is more attractive when $N = Z$.

More detailed examination reveals greater stability when $N$ and $Z$ are even numbers—nuclear forces are more attractive when neutrons and protons are in pairs. For increasingly higher masses, there are progressively more neutrons than protons in stable nuclei. This is due to the ever-growing repulsion between protons. Since nuclear forces are short ranged, and the Coulomb force is long ranged, an excess of neutrons keeps the protons a little farther apart, reducing Coulomb repulsion. Decay modes of nuclides out of the region of stability consistently produce nuclides closer to the region of stability. There are more stable nuclei having certain numbers of protons and neutrons, called magic numbers. Magic numbers indicate a shell structure for the nucleus in which closed shells are more stable. Nuclear shell theory has been very successful in explaining nuclear energy levels, nuclear decay, and the greater stability of nuclei with closed shells. We have been producing ever-heavier transuranic elements since the early 1940s, and we have now produced the element with $Z = 118$. There are theoretical predictions of an island of relative stability for nuclei with such high $Z$'s.

Figure 31.15 The German-born American physicist Maria Goeppert Mayer (1906–1972) shared the 1963 Nobel Prize in physics with J. Jensen for the creation of the nuclear shell model. This successful nuclear model has nucleons filling shells analogous to electron shells in atoms. It was inspired by patterns observed in nuclear properties. (credit: Nobel Foundation via Wikimedia Commons)

### 31.4 Nuclear Decay and Conservation Laws

Nuclear decay has provided an amazing window into the realm of the very small. Nuclear decay gave the first indication of the connection between mass and energy, and it revealed the existence of two of the four basic forces in nature. In this section, we explore the major modes of nuclear decay; and, like those who first explored them, we will discover evidence of previously unknown particles and conservation laws.
Some nuclides are stable, apparently living forever. Unstable nuclides decay (that is, they are radioactive), eventually producing a stable nuclide after many decays. We call the original nuclide the parent and its decay products the daughters. Some radioactive nuclides decay in a single step to a stable nucleus. For example, $^{60}\text{Co}$ is unstable and decays directly to $^{60}\text{Ni}$, which is stable. Others, such as $^{238}\text{U}$, decay to another unstable nuclide, resulting in a decay series in which each subsequent nuclide decays until a stable nuclide is finally produced. The decay series that starts from $^{238}\text{U}$ is of particular interest, since it produces the radioactive isotopes $^{226}\text{Ra}$ and $^{210}\text{Po}$, which the Curies first discovered (see Figure 31.16). Radon gas is also produced ($^{222}\text{Rn}$ in the series), an increasingly recognized naturally occurring hazard. Since radon is a noble gas, it emanates from materials, such as soil, containing even trace amounts of $^{238}\text{U}$ and can be inhaled. The decay of radon and its daughters produces internal damage. The $^{238}\text{U}$ decay series ends with $^{206}\text{Pb}$, a stable isotope of lead.

Figure 31.16 The decay series produced by $^{238}\text{U}$, the most common uranium isotope. Nuclides are graphed in the same manner as in the chart of nuclides. The type of decay for each member of the series is shown, as well as the half-lives. Note that some nuclides decay by more than one mode. You can see why radium and polonium are found in uranium ore. A stable isotope of lead is the end product of the series.

Note that the daughters of $\alpha$ decay shown in Figure 31.16 always have two fewer protons and two fewer neutrons than the parent. This seems reasonable, since we know that $\alpha$ decay is the emission of a $^4\text{He}$ nucleus, which has two protons and two neutrons. The daughters of $\beta$ decay have one less neutron and one more proton than their parent. Beta decay is a little more subtle, as we shall see. No $\gamma$ decays are shown in the figure, because they do not produce a daughter that differs from the parent.

Alpha Decay

In alpha decay, a $^4\text{He}$ nucleus simply breaks away from the parent nucleus, leaving a daughter with two fewer protons and two fewer neutrons than the parent (see Figure 31.17). One example of $\alpha$ decay is shown in Figure 31.16 for $^{238}\text{U}$. Another nuclide that undergoes $\alpha$ decay is $^{239}\text{Pu}$. The decay equations for these two nuclides are

$$^{238}\text{U} \rightarrow 234\text{Th} + 4\text{He} \quad (31.13)$$
and

\[ ^{239}\text{Pu} \rightarrow ^{235}\text{U} + ^{4}\text{He}. \]  

(31.14)

Figure 31.17 Alpha decay is the separation of a \(^4\text{He}\) nucleus from the parent. The daughter nucleus has two fewer protons and two fewer neutrons than the parent. Alpha decay occurs spontaneously only if the daughter and \(^4\text{He}\) nucleus have less total mass than the parent.

If you examine the periodic table of the elements, you will find that Th has \(Z = 90\), two fewer than U, which has \(Z = 92\). Similarly, in the second decay equation, we see that U has two fewer protons than Pu, which has \(Z = 94\). The general rule for \(\alpha\) decay is best written in the format

\[ \frac{A}{Z}X_N \rightarrow \frac{A-4}{Z-2}Y_{N-2} + \frac{4}{2}\text{He}_2 \]  

(31.15)

where \(Y\) is the nuclide that has two fewer protons than \(X\), such as Th having two fewer than U. So if you were told that \(^{239}\text{Pu}\) \(\alpha\) decays and were asked to write the complete decay equation, you would first look up which element has two fewer protons (an atomic number two lower) and find that this is uranium. Then since four nucleons have broken away from the original 239, its atomic mass would be 235.

It is instructive to examine conservation laws related to \(\alpha\) decay. You can see from the equation

\[ \frac{A}{Z}X_N \rightarrow \frac{A-4}{Z-2}Y_{N-2} + \frac{4}{2}\text{He}_2 \]  

that total charge is conserved. Linear and angular momentum are conserved, too. Although conserved angular momentum is not of great consequence in this type of decay, conservation of linear momentum has interesting consequences. If the nucleus is at rest when it decays, its momentum is zero. In that case, the fragments must fly in opposite directions with equal-magnitude momenta so that total momentum remains zero. This results in the \(\alpha\) particle carrying away most of the energy, as a bullet from a heavy rifle carries away most of the energy of the powder burned to shoot it. Total mass-energy is also conserved: the energy produced in the decay comes from conversion of a fraction of the original mass. As discussed in Atomic Physics, the general relationship is

\[ E = (\Delta m)c^2. \]  

(31.16)

Here, \(E\) is the nuclear reaction energy (the reaction can be nuclear decay or any other reaction), and \(\Delta m\) is the difference in mass between initial and final products. When the final products have less total mass, \(\Delta m\) is positive, and the reaction releases energy (is exothermic). When the products have greater total mass, the reaction is endothermic (\(\Delta m\) is negative) and must be induced with an energy input. For \(\alpha\) decay to be spontaneous, the decay products must have smaller mass than the parent.

**Example 31.2 Alpha Decay Energy Found from Nuclear Masses**

Find the energy emitted in the \(\alpha\) decay of \(^{239}\text{Pu}\).

**Strategy**

Nuclear reaction energy, such as released in \(\alpha\) decay, can be found using the equation \(E = (\Delta m)c^2\). We must first find \(\Delta m\), the difference in mass between the parent nucleus and the products of the decay. This is easily done using masses given in Appendix A.

**Solution**

The decay equation was given earlier for \(^{239}\text{Pu}\); it is

\[ ^{239}\text{Pu} \rightarrow ^{235}\text{U} + ^{4}\text{He}. \]  

(31.17)

Thus the pertinent masses are those of \(^{239}\text{Pu}\), \(^{235}\text{U}\), and the \(\alpha\) particle or \(^4\text{He}\), all of which are listed in Appendix A. The initial mass was \(m(^{239}\text{Pu}) = 239.052157\) u. The final mass is the sum \(m(^{235}\text{U}) + m(^4\text{He})= 235.043924\) u + 4.002602 u = 239.046526 u. Thus,

\[ \Delta m = m(^{239}\text{Pu}) - [m(^{235}\text{U}) + m(^4\text{He})] \]  

\[ = 239.052157\) u - 239.046526 u \]  

\[ = 0.005631\) u. \]  

(31.18)

Now we can find \(E\) by entering \(\Delta m\) into the equation:

\[ E = (\Delta m)c^2 = (0.005631\text{ u})c^2. \]  

(31.19)

We know 1 u = 931.5 MeV/c\(^2\), and so
\[ E = (0.005631)(931.5 \text{ MeV} / c^2)(c^2) = 5.25 \text{ MeV}. \]  

**Discussion**

The energy released in this \( \alpha \) decay is in the MeV range, about \( 10^6 \) times as great as typical chemical reaction energies, consistent with many previous discussions. Most of this energy becomes kinetic energy of the \( \alpha \) particle (or \( ^4 \text{He} \) nucleus), which moves away at high speed.

The energy carried away by the recoil of the \( ^{235} \text{U} \) nucleus is much smaller in order to conserve momentum. The \( ^{235} \text{U} \) nucleus can be left in an excited state to later emit photons (\( \gamma \) rays). This decay is spontaneous and releases energy, because the products have less mass than the parent nucleus. The question of why the products have less mass will be discussed in Binding Energy. Note that the masses given in Appendix A are atomic masses of neutral atoms, including their electrons. The mass of the electrons is the same before and after \( \alpha \) decay, and so their masses subtract out when finding \( \Delta m \). In this case, there are 94 electrons before and after the decay.

**Beta Decay**

There are actually three types of beta decay. The first discovered was “ordinary” beta decay and is called \( \beta^- \) decay or electron emission. The symbol \( \beta^- \) represents an electron emitted in nuclear beta decay. Cobalt-60 is a nuclide that \( \beta^- \) decays in the following manner:

\[ ^{60}\text{Co} \rightarrow ^{60}\text{Ni} + \beta^- + \text{neutrino}. \]  

The neutrino is a particle emitted in beta decay that was unanticipated and is of fundamental importance. The neutrino was not even proposed in theory until more than 20 years after beta decay was known to involve electron emissions. Neutrinos are so difficult to detect that the first direct evidence of them was not obtained until 1953. Neutrinos are nearly massless, have no charge, and do not interact with nucleons via the strong nuclear force. Traveling approximately at the speed of light, they have little time to affect any nucleus they encounter. This is, owing to the fact that they have no charge (and they are not EM waves), they do not interact through the EM force. They do interact via the relatively weak and very short range weak nuclear force. Consequently, neutrinos escape almost any detector and penetrate almost any shielding. However, neutrinos do carry energy, angular momentum (they are fermions with half-integral spin), and linear momentum away from a beta decay. When accurate measurements of beta decay were made, it became apparent that energy, angular momentum, and linear momentum were not accounted for by the daughter nucleus and electron alone. Either a previously unsuspected particle was carrying them away, or three conservation laws were being violated. Wolfgang Pauli made a formal proposal for the existence of neutrinos in 1930. The Italian-born American physicist Enrico Fermi (1901–1954) gave neutrinos their name, meaning little neutral ones, when he developed a sophisticated theory of beta decay (see Figure 31.18). Part of Fermi’s theory was the identification of the weak nuclear force as being distinct from the strong nuclear force and in fact responsible for beta decay.

![Figure 31.18](image_url) Enrico Fermi was nearly unique among 20th-century physicists—he made significant contributions both as an experimentalist and a theorist. His many contributions to theoretical physics included the identification of the weak nuclear force. The fermi (fm) is named after him, as are an entire class of subatomic particles (fermions), an element (Fermium), and a major research laboratory (Fermilab). His experimental work included studies of radioactivity, for which he won the 1938 Nobel Prize in physics, and creation of the first nuclear chain reaction. (credit: United States Department of Energy, Office of Public Affairs)

The neutrino also reveals a new conservation law. There are various families of particles, one of which is the electron family. We propose that the number of members of the electron family is constant in any process or any closed system. In our example of beta decay, there are no members of the electron family present before the decay, but after, there is an electron and a neutrino. So electrons are given an electron family number of \( +1 \).

The neutrino in \( \beta^- \) decay is an electron’s antineutrino, given the symbol \( \bar{\nu}_e \), where \( \nu \) is the Greek letter nu, and the subscript \( e \) means this neutrino is related to the electron. The bar indicates this is a particle of antimatter. (All particles have antimatter counterparts that are nearly identical except that they have the opposite charge. Antimatter is almost entirely absent on Earth, but it is found in nuclear decay and other nuclear and particle reactions as well as in outer space.) The electron’s antineutrino \( \bar{\nu}_e \), being antimatter, has an electron family number of \( -1 \). The total is zero, before and after the decay. The new conservation law, obeyed in all circumstances, states that the total electron family number is constant. An electron cannot be created without also creating an antimatter family member. This law is analogous to the conservation of charge in a situation where total charge is originally zero, and equal amounts of positive and negative charge must be created in a reaction to keep the total zero.
If a nuclide $^A_ZX_N$ is known to $\beta^-$ decay, then its $\beta^-$ decay equation is

$$X_N \rightarrow Y_{N-1} + \beta^- + \bar{\nu}_e \quad (\beta^- \text{ decay}), \quad (31.22)$$

where $Y$ is the nuclide having one more proton than $X$ (see Figure 31.19). So if you know that a certain nuclide $\beta^-$ decays, you can find the daughter nucleus by first looking up $Z$ for the parent and then determining which element has atomic number $Z + 1$. In the example of the $\beta^-$ decay of $^{60}\text{Co}$ given earlier, we see that $Z = 27$ for Co and $Z = 28$ is Ni. It is as if one of the neutrons in the parent nucleus decays into a proton, electron, and neutrino. In fact, neutrons outside of nuclei do just that—they live only an average of a few minutes and $\beta^-$ decay in the following manner:

$$n \rightarrow p + \beta^- + \bar{\nu}_e. \quad (31.23)$$

![Figure 31.19](image)

Figure 31.19 In $\beta^-$ decay, the parent nucleus emits an electron and an antineutrino. The daughter nucleus has one more proton and one less neutron than its parent. Neutrinos interact so weakly that they are almost never directly observed, but they play a fundamental role in particle physics.

We see that charge is conserved in $\beta^-$ decay, since the total charge is $Z$ before and after the decay. For example, in $^{60}\text{Co}$ decay, total charge is 27 before decay, since cobalt has $Z = 27$. After decay, the daughter nucleus is Ni, which has $Z = 28$, and there is an electron, so that the total charge is also $28 + (-1)$ or 27. Angular momentum is conserved, but not obviously (you have to examine the spins and angular momenta of the final products in detail to verify this). Linear momentum is also conserved, again imparting most of the decay energy to the electron and the antineutrino, since they are of low and zero mass, respectively. Another new conservation law is obeyed here and elsewhere in nature. The total number of nucleons $A$ is conserved. In $^{60}\text{Co}$ decay, for example, there are 60 nucleons before and after the decay. Note that total $A$ is also conserved in $\alpha$ decay. Also note that the total number of protons changes, as does the total number of neutrons, so that total $Z$ and total $N$ are not conserved in $\beta^-$ decay, as they are in $\alpha$ decay. Energy released in $\beta^-$ decay can be calculated given the masses of the parent and products.

Example 31.3 $\beta^-$ Decay Energy from Masses

Find the energy emitted in the $\beta^-$ decay of $^{60}\text{Co}$.

**Strategy and Concept**

As in the preceding example, we must first find $\Delta m$, the difference in mass between the parent nucleus and the products of the decay, using masses given in Appendix A. Then the emitted energy is calculated as before, using $E = (\Delta m)c^2$. The initial mass is just that of the parent nucleus, and the final mass is that of the daughter nucleus and the electron created in the decay. The neutrino is massless, or nearly so. However, since the masses given in Appendix A are for neutral atoms, the daughter nucleus has one more electron than the parent, and so the extra electron mass that corresponds to the $\beta^-$ is included in the atomic mass of Ni. Thus,

$$\Delta m = m(^{60}\text{Co}) - m(^{60}\text{Ni}). \quad (31.24)$$

**Solution**

The $\beta^-$ decay equation for $^{60}\text{Co}$ is

$$^{27}_{27}\text{Co}_{33} \rightarrow ^{28}_{28}\text{Ni}_{32} + \beta^- + \bar{\nu}_e. \quad (31.25)$$

As noticed,

$$\Delta m = m(^{60}\text{Co}) - m(^{60}\text{Ni}). \quad (31.26)$$

Entering the masses found in Appendix A gives

$$\Delta m = 59.933820 \text{ u} - 59.930789 \text{ u} = 0.003031 \text{ u}. \quad (31.27)$$

Thus,

$$E = (\Delta m)c^2 = (0.003031 \text{ u})c^2. \quad (31.28)$$

Using $1 \text{ u} = 931.5 \text{ MeV} / c^2$, we obtain
Beyond $Z$ is known to undergo electron capture, then its decay, it is as if one of the protons in the parent nucleus decays into a neutron, a positron, and a neutrino. Protons do not do this outside of decay. Electron capture is sometimes denoted by the letters EC. We know that electrons cannot reside in the nucleus, but this is $-1$ e, conserves electron family number as well as all other. When a $Z$, but in beta decay it is written as $\beta^+$ decays, the daughter nucleus $^{60}\text{Ni}$ is left in an excited state and emits photons ($\gamma$ rays). Most of the remaining energy goes to the electron and neutrino, since the recoil kinetic energy of the daughter nucleus is small. One final note: the electron emitted in $\beta^-$ decay is created in the nucleus at the time of decay.

The second type of beta decay is less common than the first. It is $\beta^+$ decay. Certain nuclides decay by the emission of a positive electron. This is antielectron or positron decay (see Figure 31.20).

![Image](image.png)

\textbf{Figure 31.20} $\beta^+$ decay is the emission of a positron that eventually finds an electron to annihilate, characteristically producing gammas in opposite directions.

The antielectron is often represented by the symbol $e^+$, but in beta decay it is written as $\beta^+$ to indicate the antielectron was emitted in a nuclear decay. Antielectrons are the antimatter counterpart to electrons, being nearly identical, having the same mass, spin, and so on, but having a positive charge and an electron family number of $-1$. When a positron encounters an electron, there is a mutual annihilation in which all the mass of the antielectron-electron pair is converted into pure photon energy. (The reaction, $e^+ + e^- \rightarrow \gamma + \gamma$, conserves electron family number as well as all other conserved quantities.) If a nuclide $\frac{A}{Z} X_N$ is known to $\beta^+$ decay, then its $\beta^+$ decay equation is

$$\frac{A}{Z} X_N \rightarrow Y_{N+1} + \beta^+ + \nu_e (\beta^+ \text{ decay}),$$

where $Y$ is the nuclide having one less proton than $X$ (to conserve charge) and $\nu_e$ is the symbol for the electron's neutrino, which has an electron family number of $+1$. Since an antimatter member of the electron family (the $\beta^+$) is created in the decay, a matter member of the family (here the $\nu_e$) must also be created. Given, for example, that $^{22}\text{Na}$ $\beta^+$ decays, you can write its full decay equation by first finding that $Z = 11$ for $^{22}\text{Na}$, so that the daughter nuclide will have $Z = 10$, the atomic number for neon. Thus the $\beta^+$ decay equation for $^{22}\text{Na}$ is

$$^{22}_{11}\text{Na}_{11} \rightarrow ^{22}_{10}\text{Ne}_{12} + \beta^+ + \nu_e.$$  

In $\beta^+$ decay, it is as if one of the protons in the parent nucleus decays into a neutron, a positron, and a neutrino. Protons do not do this outside of the nucleus, and so the decay is due to the complexities of the nuclear force. Note again that the total number of nucleons is constant in this and any other reaction. To find the energy emitted in $\beta^+$ decay, you must again count the number of electrons in the neutral atoms, since atomic masses are used. The daughter has one less electron than the parent, and one electron mass is created in the decay. Thus, in $\beta^+$ decay,

$$\Delta m = m(\text{parent}) - [m(\text{daughter}) + 2m_e],$$

since we use the masses of neutral atoms.

\textbf{Electron capture} is the third type of beta decay. Here, a nucleus captures an inner-shell electron and undergoes a nuclear reaction that has the same effect as $\beta^+$ decay. Electron capture is sometimes denoted by the letters EC. We know that electrons cannot reside in the nucleus, but this is a nuclear reaction that consumes the electron and occurs spontaneously only when the products have less mass than the parent plus the electron. If a nuclide $\frac{A}{Z} X_N$ is known to undergo electron capture, then its electron capture equation is

$$\frac{A}{Z} X_N + e^- \rightarrow Y_{N+1} + \nu_e (\text{electron capture, or EC}).$$

Any nuclide that can $\beta^+$ decay can also undergo electron capture (and often does both). The same conservation laws are obeyed for EC as for $\beta^+$ decay. It is good practice to confirm these for yourself.

All forms of beta decay occur because the parent nuclide is unstable and lies outside the region of stability in the chart of nuclides. Those nuclides that have relatively more neutrons than those in the region of stability will $\beta^-$ decay to produce a daughter with fewer neutrons, producing a...
daughter nearer the region of stability. Similarly, those nuclides having relatively more protons than those in the region of stability will $\beta^-$ decay or undergo electron capture to produce a daughter with fewer protons, nearer the region of stability.

**Gamma Decay**

Gamma decay is the simplest form of nuclear decay—it is the emission of energetic photons by nuclei left in an excited state by some earlier process. Protons and neutrons in an excited nucleus are in higher orbitals, and they fall to lower levels by photon emission (analogous to electrons in excited atoms). Nuclear excited states have lifetimes typically of only about $10^{-14}$ s, an indication of the great strength of the forces pulling the nucleons to lower states. The $\gamma$ decay equation is simply

$$\frac{A}{2} X_N^* \rightarrow X_N + \gamma_1 + \gamma_2 + \ldots \text{ (}\gamma\text{ decay)}$$

where the asterisk indicates the nucleus is in an excited state. There may be one or more $\gamma$ s emitted, depending on how the nuclide de-excites. In radioactive decay, $\gamma$ emission is common and is preceded by $\gamma$ or $\beta$ decay. For example, when $^{60}$Co $\beta^-$ decays, it most often leaves the daughter nucleus in an excited state, written $^{60}$Ni*. Then the nickel nucleus quickly $\gamma$ decays by the emission of two penetrating $\gamma$ s:

$$^{60}\text{Ni}^* \rightarrow ^{60}\text{Ni} + \gamma_1 + \gamma_2.$$  \hspace{1cm} (31.35)

These are called cobalt $\gamma$ rays, although they come from nickel—they are used for cancer therapy, for example. It is again constructive to verify the conservation laws for gamma decay. Finally, since $\gamma$ decay does not change the nuclide to another species, it is not prominently featured in charts of decay series, such as that in Figure 31.16.

There are other types of nuclear decay, but they occur less commonly than $\alpha$, $\beta$, and $\gamma$ decay. Spontaneous fission is the most important of the other forms of nuclear decay because of its applications in nuclear power and weapons. It is covered in the next chapter.

### 31.5 Half-Life and Activity

Unstable nuclei decay. However, some nuclides decay faster than others. For example, radium and polonium, discovered by the Curies, decay faster than uranium. This means they have shorter lifetimes, producing a greater rate of decay. In this section we explore half-life and activity, the quantitative terms for lifetime and rate of decay.

**Half-Life**

Why use a term like half-life rather than lifetime? The answer can be found by examining Figure 31.21, which shows how the number of radioactive nuclei in a sample decreases with time. The time in which half of the original number of nuclei decay is defined as the half-life, $t_{1/2}$. Half of the remaining nuclei decay in the next half-life. Further, half of that amount decays in the following half-life. Therefore, the number of radioactive nuclei decreases from $N$ to $N/2$ in one half-life, then to $N/4$ in the next, and to $N/8$ in the next, and so on. If $N$ is a large number, then many half-lives (not just two) pass before all of the nuclei decay. Nuclear decay is an example of a purely statistical process. A more precise definition of half-life is that each nucleus has a 50% chance of living for a time equal to one half-life $t_{1/2}$. Thus, if $N$ is reasonably large, half of the original nuclei decay in a time of one half-life. If an individual nucleus makes it through that time, it still has a 50% chance of surviving through another half-life. Even if it happens to make it through hundreds of half-lives, it still has a 50% chance of surviving through one more. The probability of decay is the same no matter when you start counting. This is like random coin flipping. The chance of heads is 50%, no matter what has happened before.

![Figure 31.21](image)

**Figure 31.21** Radioactive decay reduces the number of radioactive nuclei over time. In one half-life $t_{1/2}$, the number decreases to half of its original value. Half of what remains decay in the next half-life, and half of those in the next, and so on. This is an exponential decay, as seen in the graph of the number of nuclei present as a function of time.

There is a tremendous range in the half-lives of various nuclides, from as short as $10^{-23}$ s for the most unstable, to more than $10^{16}$ y for the least unstable, or about 46 orders of magnitude. Nuclides with the shortest half-lives are those for which the nuclear forces are least attractive, an
indication of the extent to which the nuclear force can depend on the particular combination of neutrons and protons. The concept of half-life is applicable to other subatomic particles, as will be discussed in Particle Physics. It is also applicable to the decay of excited states in atoms and nuclei. The following equation gives the quantitative relationship between the original number of nuclei present at time zero \( N_0 \) and the number \( N \) at a later time \( t \):

\[
N = N_0 e^{-\lambda t},
\]

(31.36)

where \( e = 2.71828... \) is the base of the natural logarithm, and \( \lambda \) is the decay constant for the nuclide. The shorter the half-life, the larger is the value of \( \lambda \), and the faster the exponential \( e^{-\lambda t} \) decreases with time. The relationship between the decay constant \( \lambda \) and the half-life \( t_{1/2} \) is

\[
\lambda = \frac{\ln(2)}{t_{1/2}} \approx 0.693 \frac{t}{t_{1/2}}.
\]

(31.37)

To see how the number of nuclei declines to half its original value in one half-life, let \( t = t_{1/2} \) in the exponential in the equation \( N = N_0 e^{-\lambda t} \). This gives \( N = N_0 e^{-\lambda t_{1/2}} = N_0 e^{-0.693} = 0.500N_0 \). For integral numbers of half-lives, you can just divide the original number by 2 over and over again, rather than using the exponential relationship. For example, if ten half-lives have passed, we divide \( N \) by 2 ten times. This reduces it to \( N / 1024 \). For an arbitrary time, not just a multiple of the half-life, the exponential relationship must be used.

Radioactive dating is a clever use of naturally occurring radioactivity. Its most famous application is carbon-14 dating. Carbon-14 has a half-life of 5730 years and is produced in a nuclear reaction induced when solar neutrinos strike \(^{14}\text{N}\) in the atmosphere. Radioactive carbon has the same chemistry as stable carbon, and so it mixes into the ecosphere, where it is consumed and becomes part of every living organism. Carbon-14 has an abundance of 1.3 parts per trillion of normal carbon. Thus, if you know the number of carbon nuclei in an object (perhaps determined by mass and Avogadro's number), you multiply that number by \( 1.3 \times 10^{-12} \) to find the number of \(^{14}\text{C}\) nuclei in the object. When an organism dies, carbon exchange with the environment ceases, and \(^{14}\text{C}\) is not replenished as it decays. By comparing the abundance of \(^{14}\text{C}\) in an artifact, such as mummy wrappings, with the normal abundance in living tissue, it is possible to determine the artifact's age (or time since death). Carbon-14 dating can be used for biological tissues as old as 50 or 60 thousand years, but is most accurate for younger samples, since the abundance of \(^{14}\text{C}\) nuclei in them is greater. Very old biological materials contain no \(^{14}\text{C}\) at all. There are instances in which the date of an artifact can be determined by other means, such as historical knowledge or tree-ring counting. These cross-references have confirmed the validity of carbon-14 dating and permitted us to calibrate the technique as well. Carbon-14 dating revolutionized parts of archaeology and is of such importance that it earned the 1960 Nobel Prize in chemistry for its developer, the American chemist Willard Libby (1908–1980).

One of the most famous cases of carbon-14 dating involves the Shroud of Turin, a long piece of fabric purported to be the burial shroud of Jesus (see Figure 31.22). This relic was first displayed in Turin in 1354 and was denounced as a fraud at that time by a French bishop. Its remarkable negative imprint of an apparently crucified body resembles the then-accepted image of Jesus, and so the shroud was never disregarded completely and remained controversial over the centuries. Carbon-14 dating was not performed on the shroud until 1988, when the process had been refined to the point where only a small amount of material needed to be destroyed. Samples were tested at three independent laboratories, each being given four pieces of cloth, with only one unidentified piece from the shroud, to avoid prejudice. All three laboratories found samples of the shroud contain 92% of the \(^{14}\text{C}\) found in living tissues, allowing the shroud to be dated (see Example 31.4).

Figure 31.22 Part of the Shroud of Turin, which shows a remarkable negative imprint likeness of Jesus complete with evidence of crucifixion wounds. The shroud first surfaced in the 14th century and was only recently carbon-14 dated. It has not been determined how the image was placed on the material. (credit: Butko, Wikimedia Commons)

Example 31.4 How Old Is the Shroud of Turin?

Calculate the age of the Shroud of Turin given that the amount of \(^{14}\text{C}\) found in it is 92% of that in living tissue.

Strategy
Knowing that 92% of the $^{14}\text{C}$ remains means that $N/N_0 = 0.92$. Therefore, the equation $N = N_0 e^{-\lambda t}$ can be used to find $\lambda t$. We also know that the half-life of $^{14}\text{C}$ is 5730 y, and so once $\lambda t$ is known, we can use the equation $\lambda = \frac{0.693}{t_{1/2}}$ to find $\lambda$ and then find $t$ as requested. Here, we postulate that the decrease in $^{14}\text{C}$ is solely due to nuclear decay.

**Solution**

Solving the equation $N = N_0 e^{-\lambda t}$ for $N/N_0$ gives

$$\frac{N}{N_0} = e^{-\lambda t}. \tag{31.38}$$

Thus,

$$0.92 = e^{-\lambda t}. \tag{31.39}$$

Taking the natural logarithm of both sides of the equation yields

$$\ln 0.92 = -\lambda t \tag{31.40}$$

so that

$$-0.0834 = -\lambda t. \tag{31.41}$$

Rearranging to isolate $t$ gives

$$t = \frac{0.0834}{\lambda}. \tag{31.42}$$

Now, the equation $\lambda = \frac{0.693}{t_{1/2}}$ can be used to find $\lambda$ for $^{14}\text{C}$. Solving for $\lambda$ and substituting the known half-life gives

$$\lambda = \frac{0.693}{5730 \text{ y}} = \frac{0.693}{5730 \text{ y}} \tag{31.43}$$

We enter this value into the previous equation to find $t$:

$$t = \frac{0.0834}{0.693 \text{ y}} = 690 \text{ y}. \tag{31.44}$$

**Discussion**

This dates the material in the shroud to 1988–690 = a.d. 1300. Our calculation is only accurate to two digits, so that the year is rounded to 1300. The values obtained at the three independent laboratories gave a weighted average date of a.d. 1320 ± 60. The uncertainty is typical of carbon-14 dating and is due to the small amount of $^{14}\text{C}$ in living tissues, the amount of material available, and experimental uncertainties (reduced by having three independent measurements). It is meaningful that the date of the shroud is consistent with the first record of its existence and inconsistent with the period in which Jesus lived.

There are other forms of radioactive dating. Rocks, for example, can sometimes be dated based on the decay of $^{238}\text{U}$. The decay series for $^{238}\text{U}$ ends with $^{206}\text{Pb}$, so that the ratio of these nuclides in a rock is an indication of how long it has been since the rock solidified. The original composition of the rock, such as the absence of lead, must be known with some confidence. However, as with carbon-14 dating, the technique can be verified by a consistent body of knowledge. Since $^{238}\text{U}$ has a half-life of $4.5 \times 10^9 \text{ y}$, it is useful for dating only very old materials, showing, for example, that the oldest rocks on Earth solidified about $3.5 \times 10^9$ years ago.

**Activity, the Rate of Decay**

What do we mean when we say a source is highly radioactive? Generally, this means the number of decays per unit time is very high. We define activity $R$ to be the rate of decay expressed in decays per unit time. In equation form, this is

$$R = \frac{\Delta N}{\Delta t}. \tag{31.45}$$

where $\Delta N$ is the number of decays that occur in time $\Delta t$. The SI unit for activity is one decay per second and is given the name becquerel (Bq) in honor of the discoverer of radioactivity. That is,

$$1 \text{ Bq} = 1 \text{ decay/s}. \tag{31.46}$$

Activity $R$ is often expressed in other units, such as decays per minute or decays per year. One of the most common units for activity is the curie (Ci), defined to be the activity of 1 g of $^{226}\text{Ra}$, in honor of Marie Curie’s work with radium. The definition of curie is
\[ 1 \text{ Ci} = 3.70 \times 10^{10} \text{ Bq}, \]  
(31.47)

or \(3.70 \times 10^{10}\) decays per second. A curie is a large unit of activity, while a becquerel is a relatively small unit. \(1 \text{ MBq} = 100 \text{ microcuries (} \mu \text{Ci)}\).

In countries like Australia and New Zealand that adhere more to SI units, most radioactive sources, such as those used in medical diagnostics or in physics laboratories, are labeled in Bq or megabecquerel (MBq).

Intuitively, you would expect the activity of a source to depend on two things: the amount of the radioactive substance present, and its half-life. The greater the number of radioactive nuclei present in the sample, the more will decay per unit of time. The shorter the half-life, the more decays per unit time, for a given number of nuclei. So activity \(R\) should be proportional to the number of radioactive nuclei, \(N\), and inversely proportional to their half-life, \(t_{1/2}\). In fact, your intuition is correct. It can be shown that the activity of a source is

\[ R = \frac{0.693N}{t_{1/2}}, \]  
(31.48)

where \(N\) is the number of radioactive nuclei present, having half-life \(t_{1/2}\). This relationship is useful in a variety of calculations, as the next two examples illustrate.

**Example 31.5 How Great Is the \(^{14}\text{C}\) Activity in Living Tissue?**

Calculate the activity due to \(^{14}\text{C}\) in 1.00 kg of carbon found in a living organism. Express the activity in units of Bq and Ci.

**Strategy**
To find the activity \(R\) using the equation \(R = \frac{0.693N}{t_{1/2}}\), we must know \(N\) and \(t_{1/2}\). The half-life of \(^{14}\text{C}\) can be found in Appendix B, and was stated above as 5730 y. To find \(N\), we first find the number of \(^{12}\text{C}\) nuclei in 1.00 kg of carbon using the concept of a mole. As indicated, we then multiply by \(1.3 \times 10^{-12}\) (the abundance of \(^{14}\text{C}\) in a carbon sample from a living organism) to get the number of \(^{14}\text{C}\) nuclei in a living organism.

**Solution**
One mole of carbon has a mass of 12.0 g, since it is nearly pure \(^{12}\text{C}\). (A mole has a mass in grams equal in magnitude to \(A\) found in the periodic table.) Thus the number of carbon nuclei in a kilogram is

\[ N(\text{^{12}C}) = \frac{6.02 \times 10^{23} \text{ mol}^{-1} \times (1000 \text{ g})}{12.0 \text{ g/mol}} = 5.02 \times 10^{25}. \]  
(31.49)

So the number of \(^{14}\text{C}\) nuclei in 1 kg of carbon is

\[ N(\text{^{14}C}) = (5.02 \times 10^{25})(1.3 \times 10^{-12}) = 6.52 \times 10^{13}. \]  
(31.50)

Now the activity \(R\) is found using the equation \(R = \frac{0.693N}{t_{1/2}}\).

Entering known values gives

\[ R = \frac{0.693(6.52 \times 10^{13})}{5730 \text{ y}} = 7.89 \times 10^{9} \text{ y}^{-1}, \]  
(31.51)

or \(7.89 \times 10^{9}\) decays per year. To convert this to the unit Bq, we simply convert years to seconds. Thus,

\[ R = (7.89 \times 10^{9} \text{ y}^{-1}) \frac{1.00 \text{ y}}{3.16 \times 10^{7} \text{ s}} = 250 \text{ Bq}, \]  
(31.52)

or 250 decays per second. To express \(R\) in curies, we use the definition of a curie,

\[ R = \frac{250 \text{ Bq}}{3.7 \times 10^{10} \text{ Bq/Ci}} = 6.76 \times 10^{-9} \text{ Ci}. \]  
(31.53)

Thus,

\[ R = 6.76 \text{ nCi}. \]  
(31.54)

**Discussion**
Our own bodies contain kilograms of carbon, and it is intriguing to think there are hundreds of \(^{14}\text{C}\) decays per second taking place in us. Carbon-14 and other naturally occurring radioactive substances in our bodies contribute to the background radiation we receive. The small number of decays per second found for a kilogram of carbon in this example gives you some idea of how difficult it is to detect \(^{14}\text{C}\) in a small sample of material. If there are 250 decays per second in a kilogram, then there are 0.25 decays per second in a gram of carbon in living tissue.
To observe this, you must be able to distinguish decays from other forms of radiation, in order to reduce background noise. This becomes more difficult with an old tissue sample, since it contains less $^{14}\text{C}$, and for samples more than 50 thousand years old, it is impossible.

Human-made (or artificial) radioactivity has been produced for decades and has many uses. Some of these include medical therapy for cancer, medical imaging and diagnostics, and food preservation by irradiation. Many applications as well as the biological effects of radiation are explored in *Medical Applications of Nuclear Physics*, but it is clear that radiation is hazardous. A number of tragic examples of this exist, one of the most disastrous being the meltdown and fire at the Chernobyl reactor complex in the Ukraine (see *Figure 31.23*). Several radioactive isotopes were released in huge quantities, contaminating many thousands of square kilometers and directly affecting hundreds of thousands of people. The most significant releases were of $^{131}\text{I}$, $^{90}\text{Sr}$, $^{137}\text{Cs}$, $^{239}\text{Pu}$, $^{238}\text{U}$, and $^{235}\text{U}$. Estimates are that the total amount of radiation released was about 100 million curies.

**Human and Medical Applications**

![Figure 31.23 The Chernobyl reactor. More than 100 people died soon after its meltdown, and there will be thousands of deaths from radiation-induced cancer in the future. While the accident was due to a series of human errors, the cleanup efforts were heroic. Most of the immediate fatalities were firefighters and reactor personnel. (credit: Elena Filatova)](image)

**Example 31.6 What Mass of $^{137}\text{Cs}$ Escaped Chernobyl?**

It is estimated that the Chernobyl disaster released 6.0 MCI of $^{137}\text{Cs}$ into the environment. Calculate the mass of $^{137}\text{Cs}$ released.

**Strategy**

We can calculate the mass released using Avogadro’s number and the concept of a mole if we can first find the number of nuclei $N$ released.

Since the activity $R$ is given, and the half-life of $^{137}\text{Cs}$ is found in *Appendix B* to be 30.2 y, we can use the equation $R = \frac{0.693N}{t_{1/2}}$ to find $N$.

**Solution**

Solving the equation $R = \frac{0.693N}{t_{1/2}}$ for $N$ gives

$$N = \frac{Rt_{1/2}}{0.693} \quad (31.55)$$

Entering the given values yields

$$N = \frac{(6.0 \text{ MCI})(30.2 \text{ y})}{0.693} \quad (31.56)$$

Converting curies to becquerels and years to seconds, we get

$$N = \frac{(6.0 \times 10^6 \text{ Ci})(3.7 \times 10^{10} \text{ Bq/Ci})(30.2 \text{ y})(3.16 \times 10^7 \text{ s/y})}{0.693} \quad (31.57)$$

$$= 3.1 \times 10^{26}.$$  

One mole of a nuclide $^A\text{X}$ has a mass of $A$ grams, so that one mole of $^{137}\text{Cs}$ has a mass of 137 g. A mole has $6.02 \times 10^{23}$ nuclei. Thus the mass of $^{137}\text{Cs}$ released was

$$m = \left(\frac{137 \text{ g}}{6.02 \times 10^{23}}\right)(3.1 \times 10^{26}) = 70 \times 10^3 \text{ g} \quad (31.58)$$

$$= 70 \text{ kg}.$$
Discussion

While 70 kg of material may not be a very large mass compared to the amount of fuel in a power plant, it is extremely radioactive, since it only has a 30-year half-life. Six megacuries (6.0 MCi) is an extraordinary amount of activity but is only a fraction of what is produced in nuclear reactors. Similar amounts of the other isotopes were also released at Chernobyl. Although the chances of such a disaster may have seemed small, the consequences were extremely severe, requiring greater caution than was used. More will be said about safe reactor design in the next chapter, but it should be noted that Western reactors have a fundamentally safer design.

Activity $R$ decreases in time, going to half its original value in one half-life, then to one-fourth its original value in the next half-life, and so on. Since $R = \frac{0.693N}{t^{1/2}}$, the activity decreases as the number of radioactive nuclei decreases. The equation for $R$ as a function of time is found by combining the equations $N = N_0e^{-\lambda t}$ and $R = \frac{0.693N}{t^{1/2}}$, yielding

$$R = R_0e^{-\lambda t}$$  \hspace{1cm} (31.59)

where $R_0$ is the activity at $t = 0$. This equation shows exponential decay of radioactive nuclei. For example, if a source originally has a 1.00-mCi activity, it declines to 0.500 mCi in one half-life, to 0.250 mCi in two half-lives, to 0.125 mCi in three half-lives, and so on. For times other than whole half-lives, the equation $R = R_0e^{-\lambda t}$ must be used to find $R$.

PhET Explorations: Alpha Decay

Watch alpha particles escape from a polonium nucleus, causing radioactive alpha decay. See how random decay times relate to the half life.

PhET Interactive Simulation

31.6 Binding Energy

The more tightly bound a system is, the stronger the forces that hold it together and the greater the energy required to pull it apart. We can therefore learn about nuclear forces by examining how tightly bound the nuclei are. We define the binding energy (BE) of a nucleus to be the energy required to completely disassemble it into separate protons and neutrons. We can determine the BE of a nucleus from its rest mass. The two are connected through Einstein’s famous relationship $E = (\Delta m)c^2$. A bound system has a smaller mass than its separate constituents; the more tightly the nucleons are bound together, the smaller the mass of the nucleus.

Imagine pulling a nuclide apart as illustrated in Figure 31.25. Work done to overcome the nuclear forces holding the nucleus together puts energy into the system. By definition, the energy input equals the binding energy $E$. The pieces are at rest when separated, and so the energy put into them increases their total rest mass compared with what it was when they were glued together as a nucleus. That mass increase is thus $\Delta m = \text{BE} / c^2$. This difference in mass is known as mass defect. It implies that the mass of the nucleus is less than the sum of the masses of its constituent protons and neutrons. A nuclide $^AX$ has $Z$ protons and $N$ neutrons, so that the difference in mass is

$$\Delta m = (Zm_p + Nm_n) - m_{\text{tot}}.$$  \hspace{1cm} (31.60)

Thus,

$$\text{BE} = (\Delta m)c^2 = [(Zm_p + Nm_n) - m_{\text{tot}}]c^2,$$  \hspace{1cm} (31.61)

where $m_{\text{tot}}$ is the mass of the nuclide $^AX$, $m_p$ is the mass of a proton, and $m_n$ is the mass of a neutron. Traditionally, we deal with the masses of neutral atoms. To get atomic masses into the last equation, we first add $Z$ electrons to $m_{\text{tot}}$, which gives $m(^AX)$, the atomic mass of the nuclide. We then add $Z$ electrons to the $Z$ protons, which gives $Zm(^1\text{H})$, or $Z$ times the mass of a hydrogen atom. Thus the binding energy of a nuclide $^AX$ is

$$\text{BE} = [(Zm(^1\text{H}) + Nm_n) - m(^AX)]c^2.$$  \hspace{1cm} (31.62)

The atomic masses can be found in Appendix A, most conveniently expressed in unified atomic mass units $u$ (1 $u = 931.5$ MeV / $c^2$). BE is thus calculated from known atomic masses.
Things Great and Small

Nuclear Decay Helps Explain Earth’s Hot Interior

A puzzle created by radioactive dating of rocks is resolved by radioactive heating of Earth’s interior. This intriguing story is another example of how small-scale physics can explain large-scale phenomena.

Radioactive dating plays a role in determining the approximate age of the Earth. The oldest rocks on Earth solidified about $3.5 \times 10^9$ years ago—a number determined by uranium-238 dating. These rocks could only have solidified once the surface of the Earth had cooled sufficiently. The temperature of the Earth at formation can be estimated based on gravitational potential energy of the assemblage of pieces being converted to thermal energy. Using heat transfer concepts discussed in *Thermodynamics* it is then possible to calculate how long it would take for the surface to cool to rock-formation temperatures. The result is about $10^9$ years. The first rocks formed have been solid for $3.5 \times 10^9$ years, so that the age of the Earth is approximately $4.5 \times 10^9$ years. There is a large body of other types of evidence (both Earth-bound and solar system characteristics are used) that supports this age. The puzzle is that, given its age and initial temperature, the center of the Earth should be much cooler than it is today (see Figure 31.26).

We know from seismic waves produced by earthquakes that parts of the interior of the Earth are liquid. Shear or transverse waves cannot travel through a liquid and are not transmitted through the Earth’s core. Yet compression or longitudinal waves can pass through a liquid and do go through the core. From this information, the temperature of the interior can be estimated. As noticed, the interior should have cooled more from its initial temperature in the $4.5 \times 10^9$ years since its formation. In fact, it should have taken no more than about $10^9$ years to cool to its present temperature. What is keeping it hot? The answer seems to be radioactive decay of primordial elements that were part of the material that formed the Earth (see the blowup in Figure 31.26).

Nuclides such as $^{238}$U and $^{40}$K have half-lives similar to or longer than the age of the Earth, and their decay still contributes energy to the interior. Some of the primordial radioactive nuclides have unstable decay products that also release energy—$^{238}$U has a long decay chain of these. Further, there were more of these primordial radioactive nuclides early in the life of the Earth, and thus the activity and energy contributed were greater then (perhaps by an order of magnitude). The amount of power created by these decays per cubic meter is very small. However, since a huge volume of material lies deep below the surface, this relatively small amount of energy cannot escape quickly. The power produced near the surface has much less distance to go to escape and has a negligible effect on surface temperatures.

A final effect of this trapped radiation merits mention. Alpha decay produces helium nuclei, which form helium atoms when they are stopped and capture electrons. Most of the helium on Earth is obtained from wells and is produced in this manner. Any helium in the atmosphere will escape in geologically short times because of its high thermal velocity.

What patterns and insights are gained from an examination of the binding energy of various nuclides? First, we find that BE is approximately proportional to the number of nucleons $A$ in any nucleus. About twice as much energy is needed to pull apart a nucleus like $^{24}$Mg compared with pulling apart $^{12}$C, for example. To help us look at other effects, we divide BE by $A$ and consider the binding energy per nucleon, $\text{BE}/A$. The
graph of $\text{BE}/A$ in Figure 31.27 reveals some very interesting aspects of nuclei. We see that the binding energy per nucleon averages about 8 MeV, but is lower for both the lightest and heaviest nuclei. This overall trend, in which nuclei with $A$ equal to about 60 have the greatest $\text{BE}/A$ and are thus the most tightly bound, is due to the combined characteristics of the attractive nuclear forces and the repulsive Coulomb force. It is especially important to note two things—the strong nuclear force is about 100 times stronger than the Coulomb force, and the nuclear forces are shorter in range compared to the Coulomb force. So, for low-mass nuclei, the nuclear attraction dominates and each added nucleon forms bonds with all others, causing progressively heavier nuclei to have progressively greater values of $\text{BE}/A$. This continues up to $A \approx 60$, roughly corresponding to the mass number of iron. Beyond that, new nucleons added to a nucleus will be too far from some others to feel their nuclear attraction. Added protons, however, feel the repulsion of all other protons, since the Coulomb force is longer in range. Coulomb repulsion grows for progressively heavier nuclei, but nuclear attraction remains about the same, and so $\text{BE}/A$ becomes smaller. This is why stable nuclei heavier than $A \approx 40$ have more neutrons than protons. Coulomb repulsion is reduced by having more neutrons to keep the protons farther apart (see Figure 31.28).

**Example 31.7** What Is $\text{BE}/A$ for an Alpha Particle?

Calculate the binding energy per nucleon of $^4\text{He}$, the $\alpha$ particle.

**Strategy**
To find \( \text{BE} / A \), we first find \( \text{BE} \) using the Equation: \( \text{BE} = \left[ Zm(\text{H}) + Nm_n \right] - m(A \text{X}) \) and then divide by \( A \). This is straightforward once we have looked up the appropriate atomic masses in Appendix A.

**Solution**

The binding energy for a nucleus is given by the equation

\[
\text{BE} = \left[ Zm(\text{H}) + Nm_n \right] - m(A \text{X}) c^2.
\]

For \( ^4\text{He} \), we have \( Z = N = 2 \); thus,

\[
\text{BE} = \left[ 2m(\text{H}) + 2m_n \right] - m(\text{A X}) c^2.
\]

Appendix A gives these masses as \( m(\text{He}) = 4.002602 \text{ u} \), \( m(\text{H}) = 1.007825 \text{ u} \), and \( m_n = 1.008665 \text{ u} \). Thus,

\[
\text{BE} = (0.030378 \text{ u}) c^2.
\]

Noting that \( 1 \text{ u} = 931.5 \text{ MeV}/c^2 \), we find

\[
\text{BE} = (0.030378)(931.5 \text{ MeV}/c^2)c^2 = 28.3 \text{ MeV}.
\]

Since \( A = 4 \), we see that \( \text{BE} / A \) is this number divided by 4, or

\[
\text{BE} / A = 7.07 \text{ MeV/nucleon}.
\]

**Discussion**

This is a large binding energy per nucleon compared with those for other low-mass nuclei, which have \( \text{BE} / A \approx 3 \text{ MeV/nucleon} \). This indicates that \( ^4\text{He} \) is tightly bound compared with its neighbors on the chart of the nuclides. You can see the spike representing this value of \( \text{BE} / A \) for \( ^4\text{He} \) on the graph in Figure 31.27. This is why \( ^4\text{He} \) is stable. Since \( ^4\text{He} \) is tightly bound, it has less mass than other \( A = 4 \) nuclei and, therefore, cannot spontaneously decay into them. The large binding energy also helps to explain why some nuclei undergo \( \alpha \) decay. Smaller mass in the decay products can mean energy release, and such decays can be spontaneous. Further, it can happen that two protons and two neutrons in a nucleus can randomly find themselves together, experience the exceptionally large nuclear force that binds this combination, and act as a \( ^4\text{He} \) unit within the nucleus, at least for a while. In some cases, the \( ^4\text{He} \) escapes, and \( \alpha \) decay has then taken place.

There is more to be learned from nuclear binding energies. The general trend in \( \text{BE} / A \) is fundamental to energy production in stars, and to fusion and fission energy sources on Earth, for example. This is one of the applications of nuclear physics covered in Medical Applications of Nuclear Physics. The abundance of elements on Earth, in stars, and in the universe as a whole is related to the binding energy of nuclei and has implications for the continued expansion of the universe.

**Problem-Solving Strategies**

**For Reaction And Binding Energies and Activity Calculations in Nuclear Physics**

1. **Identify exactly what needs to be determined in the problem (identify the unknowns).** This will allow you to decide whether the energy of a decay or nuclear reaction is involved, for example, or whether the problem is primarily concerned with activity (rate of decay).
2. **Make a list of what is given or can be inferred from the problem as stated (identify the knowns).**
3. **For reaction and binding-energy problems, we use atomic rather than nuclear masses.** Since the masses of neutral atoms are used, you must count the number of electrons involved. If these do not balance (such as in \( \beta^+ \) decay), then an energy adjustment of 0.511 MeV per electron must be made. Also note that atomic masses may not be given in a problem; they can be found in tables.
4. **For problems involving activity, the relationship of activity to half-life, and the number of nuclei given in the equation \( R = \frac{0.693N}{t_{1/2}} \) can be very useful.** Owing to the fact that number of nuclei is involved, you will also need to be familiar with moles and Avogadro’s number.
5. **Perform the desired calculation; keep careful track of plus and minus signs as well as powers of 10.**
6. **Check the answer to see if it is reasonable: Does it make sense?** Compare your results with worked examples and other information in the text. (Heeding the advice in Step 5 will also help you to be certain of your result.) You must understand the problem conceptually to be able to determine whether the numerical result is reasonable.

**PhET Explorations: Nuclear Fission**

Start a chain reaction, or introduce non-radioactive isotopes to prevent one. Control energy production in a nuclear reactor!
31.7 Tunneling

Protons and neutrons are bound inside nuclei, that means energy must be supplied to break them away. The situation is analogous to a marble in a bowl that can roll around but lacks the energy to get over the rim. It is bound inside the bowl (see Figure 31.30). If the marble could get over the rim, it would gain kinetic energy by rolling down outside. However classically, if the marble does not have enough kinetic energy to get over the rim, it remains forever trapped in its well.

In a nucleus, the attractive nuclear potential is analogous to the bowl at the top of a volcano (where the “volcano” refers only to the shape). Protons and neutrons have kinetic energy, but it is about 8 MeV less than that needed to get out (see Figure 31.31). That is, they are bound by an average of 8 MeV per nucleon. The slope of the hill outside the bowl is analogous to the repulsive Coulomb potential for a nucleus, such as for an α particle outside a positive nucleus. In α decay, two protons and two neutrons spontaneously break away as a 4He unit. Yet the protons and neutrons do not have enough kinetic energy to get over the rim. So how does the α particle get out?

The answer was supplied in 1928 by the Russian physicist George Gamow (1904–1968). The α particle tunnels through a region of space it is forbidden to be in, and it comes out of the side of the nucleus. Like an electron making a transition between orbits around an atom, it travels from one point to another without ever having been in between. Figure 31.32 indicates how this works. The wave function of a quantum mechanical particle varies smoothly, going from within an atomic nucleus (on one side of a potential energy barrier) to outside the nucleus (on the other side of the potential energy barrier). Inside the barrier, the wave function does not become zero but decreases exponentially, and we do not observe the particle inside the barrier. The probability of finding a particle is related to the square of its wave function, and so there is a small probability of finding the particle outside the barrier, which implies that the particle can tunnel through the barrier. This process is called barrier penetration or quantum mechanical tunneling. This concept was developed in theory by J. Robert Oppenheimer (who led the development of the first nuclear bombs during World War II) and was used by Gamow and others to describe α decay.
Figure 31.32 The wave function representing a quantum mechanical particle must vary smoothly, going from within the nucleus (to the left of the barrier) to outside the nucleus (to the right of the barrier). Inside the barrier, the wave function does not abruptly become zero; rather, it decreases exponentially. Outside the barrier, the wave function is small but finite, and there it smoothly becomes sinusoidal. Owing to the fact that there is a small probability of finding the particle outside the barrier, the particle can tunnel through the barrier.

Good ideas explain more than one thing. In addition to qualitatively explaining how the four nucleons in an $\alpha$ particle can get out of the nucleus, the detailed theory also explains quantitatively the half-life of various nuclei that undergo $\alpha$ decay. This description is what Gamow and others devised, and it works for $\alpha$ decay half-lives that vary by 17 orders of magnitude. Experiments have shown that the more energetic the $\alpha$ decay of a particular nuclide is, the shorter is its half-life. Tunneling explains this in the following manner: For the decay to be more energetic, the nucleons must have more energy in the nucleus and should be able to ascend a little closer to the rim. The barrier is therefore not as thick for more energetic decay, and the exponential decrease of the wave function inside the barrier is not as great. Thus the probability of finding the particle outside the barrier is greater, and the half-life is shorter.

Tunneling as an effect also occurs in quantum mechanical systems other than nuclei. Electrons trapped in solids can tunnel from one object to another if the barrier between the objects is thin enough. The process is the same in principle as described for $\alpha$ decay. It is far more likely for a thin barrier than a thick one. Scanning tunneling electron microscopes function on this principle. The current of electrons that travels between a probe and a sample tunnels through a barrier and is very sensitive to its thickness, allowing detection of individual atoms as shown in Figure 31.33.
activity: the rate of decay for radioactive nuclides

alpha decay: type of radioactive decay in which an atomic nucleus emits an alpha particle

alpha rays: one of the types of rays emitted from the nucleus of an atom

antinelectron: another term for positron

antimatter: composed of antiparticles

atomic mass: the total mass of the protons, neutrons, and electrons in a single atom

atomic number: number of protons in a nucleus

barrier penetration: quantum mechanical effect whereby a particle has a nonzero probability to cross through a potential energy barrier despite not having sufficient energy to pass over the barrier; also called quantum mechanical tunneling

becquerel: SI unit for rate of decay of a radioactive material

beta decay: type of radioactive decay in which an atomic nucleus emits a beta particle

beta rays: one of the types of rays emitted from the nucleus of an atom

binding energy per nucleon: the binding energy calculated per nucleon; it reveals the details of the nuclear force—larger the $\frac{BE}{A}$, the more stable the nucleus

binding energy: the energy needed to separate nucleus into individual protons and neutrons

carbon-14 dating: a radioactive dating technique based on the radioactivity of carbon-14

chart of the nuclides: a table comprising stable and unstable nuclei

curie: the activity of 1g of $^{226}$Ra, equal to $3.70 \times 10^{10}$ Bq

daughter: the nucleus obtained when parent nucleus decays and produces another nucleus following the rules and the conservation laws

decay constant: quantity that is inversely proportional to the half-life and that is used in equation for number of nuclei as a function of time

decay equation: the equation to find out how much of a radioactive material is left after a given period of time

decay series: process whereby subsequent nuclides decay until a stable nuclide is produced

decay: the process by which an atomic nucleus of an unstable atom loses mass and energy by emitting ionizing particles

electron capture equation: equation representing the electron capture

electron capture: the process in which a proton-rich nuclide absorbs an inner atomic electron and simultaneously emits a neutrino

electron's antineutrino: antiparticle of electron's neutrino

electron's neutrino: a subatomic elementary particle which has no net electric charge

Geiger tube: a very common radiation detector that usually gives an audio output

gamma decay: type of radioactive decay in which an atomic nucleus emits a gamma particle

gamma rays: one of the types of rays emitted from the nucleus of an atom

half-life: the time in which there is a 50% chance that a nucleus will decay

ionizing radiation: radiation (whether nuclear in origin or not) that produces ionization whether nuclear in origin or not
**isotopes**: nuclei having the same \( Z \) and different \( N \)’s

**magic numbers**: a number that indicates a shell structure for the nucleus in which closed shells are more stable

**mass number**: number of nucleons in a nucleus

**neutrino**: an electrically neutral, weakly interacting elementary subatomic particle

**neutron**: a neutral particle that is found in a nucleus

**nuclear radiation**: rays that originate in the nuclei of atoms, the first examples of which were discovered by Becquerel

**nuclear reaction energy**: the energy created in a nuclear reaction

**nucleons**: the particles found inside nuclei

**nucleus**: a region consisting of protons and neutrons at the center of an atom

**nuclide**: a type of atom whose nucleus has specific numbers of protons and neutrons

**parent**: the original state of nucleus before decay

**photomultiplier**: a device that converts light into electrical signals

**positron decay**: type of beta decay in which a proton is converted to a neutron, releasing a positron and a neutrino

**positron**: the particle that results from positive beta decay; also known as an antielectron

**protons**: the positively charged nucleons found in a nucleus

**quantum mechanical tunneling**: quantum mechanical effect whereby a particle has a nonzero probability to cross through a potential energy barrier despite not having sufficient energy to pass over the barrier; also called barrier penetration

**radiation detector**: a device that is used to detect and track the radiation from a radioactive reaction

**radioactive dating**: an application of radioactive decay in which the age of a material is determined by the amount of radioactivity of a particular type that occurs

**radioactive**: a substance or object that emits nuclear radiation

**radioactivity**: the emission of rays from the nuclei of atoms

**radius of a nucleus**: the radius of a nucleus is \( r = r_0 A^{1/3} \)

**range of radiation**: the distance that the radiation can travel through a material

**rate of decay**: the number of radioactive events per unit time

**scintillators**: a radiation detection method that records light produced when radiation interacts with materials

**solid-state radiation detectors**: semiconductors fabricated to directly convert incident radiation into electrical current

**tunneling**: a quantum mechanical process of potential energy barrier penetration

### Section Summary

#### 31.1 Nuclear Radioactivity
- Some nuclei are radioactive—they spontaneously decay destroying some part of their mass and emitting energetic rays, a process called nuclear radioactivity.
- Nuclear radiation, like x rays, is ionizing radiation, because energy sufficient to ionize matter is emitted in each decay.
- The range (or distance traveled in a material) of ionizing radiation is directly related to the charge of the emitted particle and its energy, with greater-charge and lower-energy particles having the shortest ranges.
- Radiation detectors are based directly or indirectly upon the ionization created by radiation, as are the effects of radiation on living and inert materials.

#### 31.2 Radiation Detection and Detectors
- Radiation detectors are based directly or indirectly upon the ionization created by radiation, as are the effects of radiation on living and inert materials.

#### 31.3 Substructure of the Nucleus
- Two particles, both called nucleons, are found inside nuclei. The two types of nucleons are protons and neutrons; they are very similar, except that the proton is positively charged while the neutron is neutral. Some of their characteristics are given in Table 31.2 and compared with those of the electron. A mass unit convenient to atomic and nuclear processes is the unified atomic mass unit (u), defined to be

\[
1 \text{ u} = 1.6605 \times 10^{-27} \text{ kg} = 931.46 \text{ MeV} / c^2.
\]
• A nuclide is a specific combination of protons and neutrons, denoted by \( ^{A}X_{N} \) or simply \( ^{A}X \),
  
  \( Z \) is the number of protons or atomic number, \( X \) is the symbol for the element, \( N \) is the number of neutrons, and \( A \) is the mass number or the total number of protons and neutrons,
  
  \[ A = N + Z. \]

• Nuclides having the same \( Z \) but different \( N \) are isotopes of the same element.

• The radius of a nucleus, \( r \), is approximately
  
  \[ r = r_{0}A^{1/3}, \]
  
  where \( r_{0} = 1.2 \text{ fm} \). Nuclear volumes are proportional to \( A \). There are two nuclear forces, the weak and the strong. Systematics in nuclear stability seen on the chart of the nuclides indicate that there are shell closures in nuclei for values of \( Z \) and \( N \) equal to the magic numbers, which correspond to highly stable nuclei.

### 31.4 Nuclear Decay and Conservation Laws

• When a parent nucleus decays, it produces a daughter nucleus following rules and conservation laws. There are three major types of nuclear decay, called alpha (\( \alpha \)), beta (\( \beta \)), and gamma (\( \gamma \)). The \( \alpha \) decay equation is
  
  \[ ^{A}X_{N} \rightarrow ^{A-4}Y_{N-2} + ^{4}\text{He}. \]

• Nuclear decay releases an amount of energy \( E \) related to the mass destroyed \( \Delta m \) by
  
  \[ E = (\Delta m)c^{2}. \]

• There are three forms of beta decay. The \( \beta^{-} \) decay equation is
  
  \[ ^{A}X_{N} \rightarrow ^{A}Y_{N-1} + \beta^{-} + \bar{\nu}_{e}. \]

• The \( \beta^{+} \) decay equation is
  
  \[ ^{A}X_{N} \rightarrow ^{A-1}Y_{N+1} + \beta^{+} + \nu_{e}. \]

• The electron capture equation is
  
  \[ ^{A}X_{N} + e^{-} \rightarrow ^{A-1}Y_{N+1} + \nu_{e}. \]

\( \beta^{-} \) is an electron, \( \beta^{+} \) is an antielectron or positron, \( \nu_{e} \) represents an electron's neutrino, and \( \bar{\nu}_{e} \) is an electron's antineutrino. In addition to all previously known conservation laws, two new ones arise—conservation of electron family number and conservation of the total number of nucleons. The \( \gamma \) decay equation is
  
  \[ X^{*}_{N} \rightarrow X_{N} + \gamma_{1} + \gamma_{2} + \cdots \]

\( \gamma \) is a high-energy photon originating in a nucleus.

### 31.5 Half-Life and Activity

• Half-life \( t_{1/2} \) is the time in which there is a 50% chance that a nucleus will decay. The number of nuclei \( N \) as a function of time is
  
  \[ N = N_{0}e^{-\lambda t}. \]

  where \( N_{0} \) is the number present at \( t = 0 \), and \( \lambda \) is the decay constant, related to the half-life by
  
  \[ \lambda = \frac{0.693}{t_{1/2}}. \]

• One of the applications of radioactive decay is radioactive dating, in which the age of a material is determined by the amount of radioactive decay that occurs. The rate of decay is called the activity \( R \):
  
  \[ R = \frac{\Delta N}{\Delta t}. \]

• The SI unit for \( R \) is the becquerel (Bq), defined by
  
  1 Bq = 1 decay/s.

• \( R \) is also expressed in terms of curies (Ci), where
  
  1 Ci = 3.70×10^{10} Bq.

• The activity \( R \) of a source is related to \( N \) and \( t_{1/2} \) by
  
  \[ R = \frac{0.693N}{t_{1/2}}. \]

• Since \( N \) has an exponential behavior as in the equation \( N = N_{0}e^{-\lambda t} \), the activity also has an exponential behavior, given by
  
  \[ R = R_{0}e^{-\lambda t}, \]

  where \( R_{0} \) is the activity at \( t = 0 \).
31.6 Binding Energy
• The binding energy (BE) of a nucleus is the energy needed to separate it into individual protons and neutrons. In terms of atomic masses,

\[ \text{BE} = \left( Z m(1\text{H}) + N m_n \right) - m(A \text{X}) c^2, \]

where \( m(1\text{H}) \) is the mass of a hydrogen atom, \( m(A \text{X}) \) is the atomic mass of the nuclide, and \( m_n \) is the mass of a neutron. Patterns in the binding energy per nucleon, \( \text{BE} / A \), reveal details of the nuclear force. The larger the \( \text{BE} / A \), the more stable the nucleus.

31.7 Tunneling
• Tunneling is a quantum mechanical process of potential energy barrier penetration. The concept was first applied to explain \( \alpha \) decay, but tunneling is found to occur in other quantum mechanical systems.

Conceptual Questions

31.1 Nuclear Radioactivity
1. Suppose the range for 5.0 MeV \( \alpha \) ray is known to be 2.0 mm in a certain material. Does this mean that every 5.0 MeV \( \alpha \) ray that strikes this material travels 2.0 mm, or does the range have an average value with some statistical fluctuations in the distances traveled? Explain.
2. What is the difference between \( \gamma \) rays and characteristic x rays? Is either necessarily more energetic than the other? Which can be the most energetic?
3. Ionizing radiation interacts with matter by scattering from electrons and nuclei in the substance. Based on the law of conservation of momentum and energy, explain why electrons tend to absorb more energy than nuclei in these interactions.
4. What characteristics of radioactivity show it to be nuclear in origin and not atomic?
5. What is the source of the energy emitted in radioactive decay? Identify an earlier conservation law, and describe how it was modified to take such processes into account.
6. Consider Figure 31.3. If an electric field is substituted for the magnetic field with positive charge instead of the north pole and negative charge instead of the south pole, in which directions will the \( \alpha \), \( \beta \), and \( \gamma \) rays bend?
7. Explain how an \( \alpha \) particle can have a larger range in air than a \( \beta \) particle with the same energy in lead.
8. Arrange the following according to their ability to act as radiation shields, with the best first and worst last. Explain your ordering in terms of how radiation loses its energy in matter.
   (a) A solid material with low density composed of low-mass atoms.
   (b) A gas composed of high-mass atoms.
   (c) A gas composed of low-mass atoms.
   (d) A solid with high density composed of high-mass atoms.
9. Often, when people have to work around radioactive materials spills, we see them wearing white coveralls (usually a plastic material). What types of radiation (if any) do you think these suits protect the worker from, and how?

31.2 Radiation Detection and Detectors
10. Is it possible for light emitted by a scintillator to be too low in frequency to be used in a photomultiplier tube? Explain.

31.3 Substructure of the Nucleus
11. The weak and strong nuclear forces are basic to the structure of matter. Why do we not experience them directly?
12. Define and make clear distinctions between the terms neutron, nucleon, nucleus, nuclide, and neutrino.
13. What are isotopes? Why do different isotopes of the same element have similar chemistries?

31.4 Nuclear Decay and Conservation Laws
14. Star Trek fans have often heard the term “antimatter drive.” Describe how you could use a magnetic field to trap antimatter, such as produced by nuclear decay, and later combine it with matter to produce energy. Be specific about the type of antimatter, the need for vacuum storage, and the fraction of matter converted into energy.
15. What conservation law requires an electron’s neutrino to be produced in electron capture? Note that the electron no longer exists after it is captured by the nucleus.
16. Neutrinos are experimentally determined to have an extremely small mass. Huge numbers of neutrinos are created in a supernova at the same time as massive amounts of light are first produced. When the 1987A supernova occurred in the Large Magellanic Cloud, visible primarily in the Southern Hemisphere and some 100,000 light-years away from Earth, neutrinos from the explosion were observed at about the same time as the light from the blast. How could the relative arrival times of neutrinos and light be used to place limits on the mass of neutrinos?
17. What do the three types of beta decay have in common that is distinctly different from alpha decay?

31.5 Half-Life and Activity
18. In a $3 \times 10^9$-year-old rock that originally contained some $^{238}\text{U}$, which has a half-life of $4.5 \times 10^9$ years, we expect to find some $^{238}\text{U}$ remaining in it. Why are $^{226}\text{Ra}$, $^{222}\text{Rn}$, and $^{210}\text{Po}$ also found in such a rock, even though they have much shorter half-lives (1600 years, 3.8 days, and 138 days, respectively)?

19. Does the number of radioactive nuclei in a sample decrease to exactly half its original value in one half-life? Explain in terms of the statistical nature of radioactive decay.

20. Radioactivity depends on the nucleus and not the atom or its chemical state. Why, then, is one kilogram of uranium more radioactive than one kilogram of uranium hexafluoride?

21. Explain how a bound system can have less mass than its components. Why is this not observed classically, say for a building made of bricks?

22. Spontaneous radioactive decay occurs only when the decay products have less mass than the parent, and it tends to produce a daughter that is more stable than the parent. Explain how this is related to the fact that more tightly bound nuclei are more stable. (Consider the binding energy per nucleon.)

23. To obtain the most precise value of BE from the equation $BE=\left[ZM(1\text{ H})+Nm\right]c^2-m(4\text{ X})c^2$, we should take into account the binding energy of the electrons in the neutral atoms. Will doing this produce a larger or smaller value for BE? Why is this effect usually negligible?

24. How does the finite range of the nuclear force relate to the fact that $BE/A$ is greatest for $A$ near 60?

31.6 Binding Energy

25. Why is the number of neutrons greater than the number of protons in stable nuclei having $A$ greater than about 40, and why is this effect more pronounced for the heaviest nuclei?

31.7 Tunneling

26. A physics student caught breaking conservation laws is imprisoned. She leans against the cell wall hoping to tunnel out quantum mechanically. Explain why her chances are negligible. (This is so in any classical situation.)

27. When a nucleus $\alpha$ decays, does the $\alpha$ particle move continuously from inside the nucleus to outside? That is, does it travel each point along an imaginary line from inside to out? Explain.
14. (a) Show that if you assume the average nucleus is spherical with a radius \( r = r_0 A^{1/3} \), and with a mass of \( A \) u, then its density is independent of \( A \).

(b) Calculate that density in u/fm\(^3\) and kg/m\(^3\), and compare your results with those found in Example 31.1 for \(^{56}\)Fe.

15. What is the ratio of the velocity of a 5.00-MeV \( \beta^- \) ray to that of an \( \alpha \) particle with the same kinetic energy? This should confirm that \( \beta^- \)s travel much faster than \( \alpha \)s even when relativity is taken into consideration. (See also Exercise 31.11.)

16. (a) What is the kinetic energy in MeV of a \( \beta^- \) ray that is traveling at 0.999c? This gives some idea of how energetic a \( \beta^- \) ray must be to travel at nearly the same speed as a \( \gamma \) ray. (b) What is the velocity of the \( \gamma \) ray relative to the \( \beta^- \) ray?

### 31.4 Nuclear Decay and Conservation Laws

In the following eight problems, write the complete decay equation for the given nuclide in the complete \( \frac{A}{2}X_{N'} \) notation. Refer to the periodic table for values of \( Z \).

17. \( \beta^- \) decay of \(^3\)H (tritium), a manufactured isotope of hydrogen used in some digital watch displays, and manufactured primarily for use in hydrogen bombs.

18. \( \beta^- \) decay of \(^{40}\)K, a naturally occurring rare isotope of potassium responsible for some of our exposure to background radiation.

19. \( \beta^+ \) decay of \(^{50}\)Mn.

20. \( \beta^+ \) decay of \(^{52}\)Fe.

21. Electron capture by \(^7\)Be.

22. Electron capture by \(^{106}\)Sn.

23. \( \alpha \) decay of \(^{210}\)Po, the isotope of polonium in the decay series of \(^{238}\)U that was discovered by the Curie family. A favorite isotope in physics labs, since it has a short half-life and decays to a stable nuclide.

24. \( \alpha \) decay of \(^{226}\)Ra, another isotope in the decay series of \(^{238}\)U, first recognized as a new element by the Curie family. Poses special problems because its daughter is a radioactive noble gas.

In the following four problems, identify the parent nuclide and write the complete decay equation in the \( \frac{A}{2}X_{N'} \) notation. Refer to the periodic table for values of \( Z \).

25. \( \beta^- \) decay producing \(^{137}\)Ba. The parent nuclide is a major waste product of reactors and has chemistry similar to potassium and sodium, resulting in its concentration in your cells if ingested.

26. \( \beta^- \) decay producing \(^{90}\)Y. The parent nuclide is a major waste product of reactors and has chemistry similar to calcium, so that it is concentrated in bones if ingested (\(^{90}\)Y is also radioactive.)

27. \( \alpha \) decay producing \(^{228}\)Ra. The parent nuclide is nearly 100% of the natural element and is found in gas lantern mantles and in metal alloys used in jets (\(^{228}\)Ra is also radioactive).
28. $\alpha$ decay producing $^{208}$Pb. The parent nuclide is in the decay series produced by $^{232}$Th, the only naturally occurring isotope of thorium.  

29. When an electron and positron annihilate, both their masses are destroyed, creating two equal energy photons to preserve momentum. (a) Confirm that the annihilation equation $e^+ + e^- \rightarrow \gamma + \gamma$ conserves charge, electron family number, and total number of nucleons. To do this, identify the values of each before and after the annihilation. (b) Find the energy of each $\gamma$ ray, assuming the electron and positron are initially nearly at rest. (c) Explain why the two $\gamma$ rays travel in exactly opposite directions if the center of mass of the electron-positron system is initially at rest.  

30. Confirm that charge, electron family number, and the total number of nucleons are all conserved by the rule for $\alpha$ decay given in the equation $^{A}Z\;X_{N} \rightarrow ^{A-4}Z-2\;Y_{N-2} + ^{4}\text{He}_2$. To do this, identify the values of each before and after the decay.  

31. Confirm that charge, electron family number, and the total number of nucleons are all conserved by the rule for $\beta^-$ decay given in the equation $^{A}Z\;X_{N} \rightarrow ^{A}\;Y_{N-1} + \beta^- + \gamma$. To do this, identify the values of each before and after the decay.  

32. Confirm that charge, electron family number, and the total number of nucleons are all conserved by the rule for $\beta^+$ decay given in the equation $^{A}Z\;X_{N} \rightarrow ^{A}Z-1\;Y_{N-1} + \beta^+ + \gamma$. To do this, identify the values of each before and after the decay.  

33. Confirm that charge, electron family number, and the total number of nucleons are all conserved by the rule for electron capture given in the equation $^{A}Z\;X_{N} + e^- \rightarrow ^{A}Z-1\;Y_{N+1} + \gamma$. To do this, identify the values of each before and after the capture.  

34. A rare decay mode has been observed in which $^{222}$Ra emits a $^{14}$C nucleus. (a) The decay equation is $^{222}$Ra $\rightarrow ^{4}$X$+^{14}$C. Identify the nuclide $^{A}$X. (b) Find the energy emitted in the decay. The mass of $^{222}$Ra is 222.015353 u.  

35. (a) Write the complete $\alpha$ decay equation for $^{222}$Ra. (b) Find the energy released in the decay.  

36. (a) Write the complete $\alpha$ decay equation for $^{249}$Cf. (b) Find the energy released in the decay.  

37. (a) Write the complete $\beta^-$ decay equation for the neutron. (b) Find the energy released in the decay.  

38. (a) Write the complete $\beta^+$ decay equation for $^{90}$Sr, a major waste product of nuclear reactors. (b) Find the energy released in the decay.  

39. Calculate the energy released in the $\beta^+$ decay of $^{22}$Na, the equation for which is given in the text. The masses of $^{22}$Na and $^{22}$Ne are 21.994434 and 21.991383 u, respectively.  

40. (a) Write the complete $\beta^+$ decay equation for $^{11}$C. (b) Calculate the energy released in the decay. The masses of $^{11}$C and $^{11}$B are 11.011433 and 11.009305 u, respectively.  

41. (a) Calculate the energy released in the $\alpha$ decay of $^{238}$U. (b) What fraction of the mass of a single $^{238}$U is destroyed in the decay? The mass of $^{234}$Th is 234.043593 u. (c) Although the fractional mass loss is large for a single nucleus, it is difficult to observe for an entire macroscopic sample of uranium. Why is this?  

42. (a) Write the complete reaction equation for electron capture by $^{7}$Be. (b) Calculate the energy released.  

43. (a) Write the complete reaction equation for electron capture by $^{15}$O. (b) Calculate the energy released.  

31.5 Half-Life and Activity  

Data from the appendices and the periodic table may be needed for these problems.  

44. An old campfire is uncovered during an archaeological dig. Its charcoal is found to contain less than 1/1000 the normal amount of $^{14}$C. Estimate the minimum age of the charcoal, noting that $2^{10} = 1024$.  

45. A $^{60}$Co source is labeled 4.00 mCi, but its present activity is found to be $1.85 \times 10^7$ Bq. (a) What is the present activity in mCi? (b) How long ago did it actually have a 4.00-mCi activity?  

46. (a) Calculate the activity $R$ in curies of 1.00 g of $^{226}$Ra. (b) Discuss why your answer is not exactly 1.00 Ci, given that the curie was originally supposed to be exactly the activity of a gram of radium.  

47. Show that the activity of the $^{14}$C in 1.00 g of $^{12}$C found in living tissue is 0.250 Bq.  

48. Mantles for gas lanterns contain thorium, because it forms an oxide that can survive being heated to incandescence for long periods of time. Natural thorium is almost 100% $^{232}$Th, with a half-life of 1.405 $\times 10^{10}$ y. If an average lantern mantle contains 300 mg of thorium, what is its activity?  

49. Cow's milk produced near nuclear reactors can be tested for as little as 1.00 pCi of $^{131}$I per liter, to check for possible reactor leakage. What mass of $^{131}$I has this activity?  

50. (a) Natural potassium contains $^{40}$K, which has a half-life of 1.277 $\times 10^9$ y. What mass of $^{40}$K in a person would have a decay rate of 4140 Bq? (b) What is the fraction of $^{40}$K in natural potassium, given that the person has 140 g in his body? (These numbers are typical for a 70-kg adult.)  

51. There is more than one isotope of natural uranium. If a researcher isolates 1.00 mg of the relatively scarce $^{235}$U and finds this mass to have an activity of 80.0 Bq, what is its half-life in years?  

52. $^{50}$V has one of the longest known radioactive half-lives. In a difficult experiment, a researcher found that the activity of 1.00 kg of $^{50}$V is 1.75 Bq. What is the half-life in years?  

53. You can sometimes find deep red crystal vases in antique stores, called uranium glass because their color was produced by doping the glass with uranium. Look up the natural isotopes of uranium and their
54. A tree falls in a forest. How many years must pass before the $^{14}$C activity in 1.00 g of the tree’s carbon drops to 1.00 decay per hour?

55. What fraction of the $^{40}$K that was on Earth when it formed 4.5x10^9 years ago is left today?

56. A 5000-Ci $^{60}$Co source used for cancer therapy is considered too weak to be useful when its activity falls to 3500 Ci. How long after its manufacture does this happen?

57. Natural uranium is 0.7200% $^{235}$U and 99.27% $^{238}$U. What were the percentages of $^{235}$U and $^{238}$U in natural uranium when Earth formed 4.5x10^9 years ago?

58. The $\beta^-$ particles emitted in the decay of $^3$H (tritium) interact with matter to create light in a glow-in-the-dark exit sign. At the time of manufacture, such a sign contains 15.0 Ci of $^3$H. (a) What is the mass of the tritium? (b) What is its activity? (c) What is its half-life after manufacture?

59. World War II aircraft had instruments with glowing radium-painted dials (see Figure 31.2). The activity of one such instrument was 1.0x10^5 Bq when new. (a) What mass of $^{226}$Ra was present? (b) After some years, the phosphors on the dials deteriorated chemically, but the radium did not escape. What is the activity of this instrument 57.0 years after it was made?

60. (a) The $^{210}$Po source used in a physics laboratory is labeled as having an activity of 1.0 $\mu$Ci on the date it was prepared. A student measures the radioactivity of this source with a Geiger counter and observes 1500 counts per minute. She notices that the source was prepared 120 days before her lab. What fraction of the decays is she observing with her apparatus? (b) Identify some of the reasons that only a fraction of the $\alpha$'s emitted are observed by the detector.

61. Armor-piercing shells with depleted uranium cores are fired by aircraft at tanks. (The high density of the uranium makes them effective.) The uranium is called depleted because it has had its $^{235}$U removed for reactor use and is nearly pure $^{238}$U. Depleted uranium has been erroneously called non-radioactive. To demonstrate that this is wrong: (a) Calculate the activity of 60.0 g of pure $^{238}$U. (b) Calculate the activity of 60.0 g of natural uranium, neglecting the $^{234}$U and all daughter nuclides.

62. The ceramic glaze on a red-orange Fiestaware plate is $U_2O_3$ and contains 50.0 grams of $^{238}$U, but very little $^{235}$U. (a) What is the activity of the plate? (b) Calculate the total energy that will be released by the $^{238}$U decay. (c) If energy is worth 12.0 cents per kW-h, what is the monetary value of the energy emitted? (These plates went out of production some 30 years ago, but are still available as collectibles.)

63. Large amounts of depleted uranium ($^{238}$U) are available as a by-product of uranium processing for reactor fuel and weapons. Uranium is very dense and makes good counter weights for aircraft. Suppose you have a 4000-kg block of $^{238}$U. (a) Find its activity. (b) How many calories per day are generated by thermalization of the decay energy? (c) Do you think you could detect this as heat? Explain.

64. The Galileo space probe was launched on its long journey past several planets in 1989, with an ultimate goal of Jupiter. Its power source is 11.0 kg of $^{238}$Pu, a by-product of nuclear weapons plutonium production. Electrical energy is generated thermoelctrically from the heat produced when the 5.59-MeV $\alpha$ particles emitted in each decay crash to a halt inside the plutonium and its shielding. The half-life of $^{238}$Pu is 87.7 years. (a) What was the original activity of the $^{238}$Pu in becquerels? (b) What power was emitted in kilowatts? (c) What power was emitted 12.0 y after launch? You may neglect any extra energy from daughter nuclides and any losses from escaping $\gamma$ rays.

65. Construct Your Own Problem

Consider the generation of electricity by a radioactive isotope in a space probe, such as described in Exercise 31.64. Construct a problem in which you calculate the mass of a radioactive isotope you need in order to supply power for a long space flight. Among the things to consider are the isotope chosen, its half-life and decay energy, the power needs of the probe and the length of the flight.

66. Unreasonable Results

A nuclear physicist finds 1.0 $\mu$g of $^{236}$U in a piece of uranium ore and assumes it is primordial since its half-life is $2.3\times10^7$ y. (a) Calculate the amount of $^{236}$U that would have had to be on Earth when it formed 4.5x10^9 y ago for 1.0 $\mu$g to be left today. (b) What is unreasonable about this result? (c) What assumption is responsible?

67. Unreasonable Results

(a) Repeat Exercise 31.57 but include the 0.0055% natural abundance of $^{234}$U with its $2.45\times10^3$ y half-life. (b) What is unreasonable about this result? (c) What assumption is responsible? (d) Where does the $^{234}$U come from if it is not primordial?

68. Unreasonable Results

The manufacturer of a smoke alarm decides that the smallest current of $\alpha$ radiation he can detect is 1.00 $\mu$A. (a) Find the activity in curies of an $\alpha$ emitter that produces a 1.00 $\mu$A current of $\alpha$ particles. (b) What is unreasonable about this result? (c) What assumption is responsible?

31.6 Binding Energy

69. $^2$H is a loosely bound isotope of hydrogen. Called deuterium or heavy hydrogen, it is stable but relatively rare—it is 0.015% of natural hydrogen. Note that deuterium has $Z = N$, which should tend to make it more tightly bound, but both are odd numbers. Calculate BE/A, the binding energy per nucleon, for $^2$H and compare it with the approximate value obtained from the graph in Figure 31.27.

70. $^{56}$Fe is among the most tightly bound of all nuclides. It is more than 90% of natural iron. Note that $^{56}$Fe has even numbers of both protons and neutrons. Calculate BE/A, the binding energy per nucleon, for $^{56}$Fe and compare it with the approximate value obtained from the graph in Figure 31.27.

71. $^{209}$Bi is the heaviest stable nuclide, and its BE/A is low compared with medium-mass nuclides. Calculate BE/A, the binding energy per nucleon, for $^{209}$Bi and compare it with the approximate value obtained from the graph in Figure 31.27.

72. (a) Calculate BE/A for $^{235}$U, the rarer of the two most common uranium isotopes. (b) Calculate BE/A for $^{238}$U. (Most of uranium is
Note that $^{238}\text{U}$ has even numbers of both protons and neutrons. Is the $\text{BE}/A$ of $^{238}\text{U}$ significantly different from that of $^{235}\text{U}$?

73. (a) Calculate $\text{BE}/A$ for $^{12}\text{C}$. Stable and relatively tightly bound, this nuclide is most of natural carbon. (b) Calculate $\text{BE}/A$ for $^{14}\text{C}$. Is the difference in $\text{BE}/A$ between $^{12}\text{C}$ and $^{14}\text{C}$ significant? One is stable and common, and the other is unstable and rare.

74. The fact that $\text{BE}/A$ is greatest for $A \approx 60$ implies that the range of the nuclear force is about the diameter of such nuclides. (a) Calculate the diameter of an $A = 60$ nucleus. (b) Compare $\text{BE}/A$ for $^{58}\text{Ni}$ and $^{90}\text{Sr}$. The first is one of the most tightly bound nuclides, while the second is larger and less tightly bound.

75. The purpose of this problem is to show in three ways that the binding energy of the electron in a hydrogen atom is negligible compared with the masses of the proton and electron. (a) Calculate the mass equivalent in u of the 13.6-eV binding energy of an electron in a hydrogen atom, and compare this with the mass of the hydrogen atom obtained from Appendix A. (b) Subtract the mass of the proton given in Table 31.2 from the mass of the hydrogen atom given in Appendix A. You will find the difference is equal to the electron’s mass to three digits, implying the binding energy is small in comparison. (c) Take the ratio of the binding energy of the electron (13.6 eV) to the energy equivalent of the electron’s mass (0.511 MeV). (d) Discuss how your answers confirm the stated purpose of this problem.

76. **Unreasonable Results**

A particle physicist discovers a neutral particle with a mass of 2.02733 u that he assumes is two neutrons bound together. (a) Find the binding energy. (b) What is unreasonable about this result? (c) What assumptions are unreasonable or inconsistent?

### 31.7 Tunneling

#### 77. Derive an approximate relationship between the energy of $\alpha$ decay and half-life using the following data. It may be useful to graph the log of $t_{1/2}$ against $E_\alpha$ to find some straight-line relationship.

<table>
<thead>
<tr>
<th>Nuclide</th>
<th>$E_\alpha$ (MeV)</th>
<th>$t_{1/2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{216}\text{Ra}$</td>
<td>9.5</td>
<td>0.18 $\mu$s</td>
</tr>
<tr>
<td>$^{194}\text{Po}$</td>
<td>7.0</td>
<td>0.7 s</td>
</tr>
<tr>
<td>$^{240}\text{Cm}$</td>
<td>6.4</td>
<td>27 d</td>
</tr>
<tr>
<td>$^{226}\text{Ra}$</td>
<td>4.91</td>
<td>1600 y</td>
</tr>
<tr>
<td>$^{232}\text{Th}$</td>
<td>4.1</td>
<td>$1.4\times10^{10}$ y</td>
</tr>
</tbody>
</table>

#### 78. Integrated Concepts

A 2.00-T magnetic field is applied perpendicular to the path of charged particles in a bubble chamber. What is the radius of curvature of the path of a 10 MeV proton in this field? Neglect any slowing along its path.

79. (a) Write the decay equation for the $\alpha$ decay of $^{235}\text{U}$. (b) What energy is released in this decay? The mass of the daughter nuclide is 231.036298 u. (c) Assuming the residual nucleus is formed in its ground state, how much energy goes to the $\alpha$ particle?

80. **Unreasonable Results**

The relatively scarce naturally occurring calcium isotope $^{48}\text{Ca}$ has a half-life of about $2\times10^{16}$ y. (a) A small sample of this isotope is labeled as having an activity of 1.0 Ci. What is the mass of the $^{48}\text{Ca}$ in the sample? (b) What is unreasonable about this result? (c) What assumption is responsible?

### 81. Unreasonable Results

A physicist scatters $\gamma$ rays from a substance and sees evidence of a nucleus $7.5\times10^{-13}$ m in radius. (a) Find the atomic mass of such a nucleus. (b) What is unreasonable about this result? (c) What is unreasonable about the assumption?

### 82. Unreasonable Results

A frazzled theoretical physicist reckons that all conservation laws are obeyed in the decay of a proton into a neutron, positron, and neutrino (as in $\beta^+$ decay of a nucleus) and sends a paper to a journal to announce the reaction as a possible end of the universe due to the spontaneous decay of protons. (a) What energy is released in this decay? (b) What is unreasonable about this result? (c) What assumption is responsible?

### 83. Construct Your Own Problem

Consider the decay of radioactive substances in the Earth’s interior. The energy emitted is converted to thermal energy that reaches the earth’s surface and is radiated away into cold dark space. Construct a problem in which you estimate the activity in a cubic meter of earth rock? And then calculate the power generated. Calculate how much power must cross each square meter of the Earth’s surface if the power is dissipated at the same rate as it is generated. Among the things to consider are the activity per cubic meter, the energy per decay, and the size of the Earth.
32.1. Medical Imaging and Diagnostics
- Explain the working principle behind an anger camera.
- Describe the SPECT and PET imaging techniques.

32.2. Biological Effects of Ionizing Radiation
- Define various units of radiation.
- Describe RBE.

32.3. Therapeutic Uses of Ionizing Radiation
- Explain the concept of radiotherapy and list typical doses for cancer therapy.

32.4. Food Irradiation
- Define food irradiation low dose, and free radicals.

32.5. Fusion
- Define nuclear fusion.
- Discuss processes to achieve practical fusion energy generation.

32.6. Fission
- Define nuclear fission.
- Discuss how fission fuel reacts and describe what it produces.
- Describe controlled and uncontrolled chain reactions.

32.7. Nuclear Weapons
- Discuss different types of fission and thermonuclear bombs.
- Explain the ill effects of nuclear explosion.
Introduction to Applications of Nuclear Physics

Applications of nuclear physics have become an integral part of modern life. From the bone scan that detects a cancer to the radioiodine treatment that cures another, nuclear radiation has diagnostic and therapeutic effects on medicine. From the fission power reactor to the hope of controlled fusion, nuclear energy is now commonplace and is a part of our plans for the future. Yet, the destructive potential of nuclear weapons haunts us, as does the possibility of nuclear reactor accidents. Certainly, several applications of nuclear physics escape our view, as seen in Figure 32.2. Not only has nuclear physics revealed secrets of nature, it has an inevitable impact based on its applications, as they are intertwined with human values. Because of its potential for alleviation of suffering, and its power as an ultimate destructor of life, nuclear physics is often viewed with ambivalence. But it provides perhaps the best example that applications can be good or evil, while knowledge itself is neither.

Figure 32.2 Customs officers inspect vehicles using neutron irradiation. Cars and trucks pass through portable x-ray machines that reveal their contents. (credit: Gerald L. Nino, CBP, U.S. Dept. of Homeland Security)

Figure 32.3 This image shows two stowaways caught illegally entering the United States from Canada. (credit: U.S. Customs and Border Protection)

32.1 Medical Imaging and Diagnostics

A host of medical imaging techniques employ nuclear radiation. What makes nuclear radiation so useful? First, $\gamma$ radiation can easily penetrate tissue; hence, it is a useful probe to monitor conditions inside the body. Second, nuclear radiation depends on the nuclide and not on the chemical compound it is in, so that a radioactive nuclide can be put into a compound designed for specific purposes. The compound is said to be tagged. A tagged compound used for medical purposes is called a radiopharmaceutical. Radiation detectors external to the body can determine the location and concentration of a radiopharmaceutical to yield medically useful information. For example, certain drugs are concentrated in inflamed regions of the body, and this information can aid diagnosis and treatment as seen in Figure 32.4. Another application utilizes a radiopharmaceutical which the body sends to bone cells, particularly those that are most active, to detect cancerous tumors or healing points. Images can then be produced of such bone scans. Radioisotopes are also used to determine the functioning of body organs, such as blood flow, heart muscle activity, and iodine uptake in the thyroid gland.

Figure 32.4 A radiopharmaceutical is used to produce this brain image of a patient with Alzheimer’s disease. Certain features are computer enhanced. (credit: National Institutes of Health)
Medical Application

Table 32.1 lists certain medical diagnostic uses of radiopharmaceuticals, including isotopes and activities that are typically administered. Many organs can be imaged with a variety of nuclear isotopes replacing a stable element by a radioactive isotope. One common diagnostic employs iodine to image the thyroid, since iodine is concentrated in that organ. The most active thyroid cells, including cancerous cells, concentrate the most iodine and, therefore, emit the most radiation. Conversely, hypothyroidism is indicated by lack of iodine uptake. Note that there is more than one isotope that can be used for several types of scans. Another common nuclear diagnostic is the thallium scan for the cardiovascular system, particularly used to evaluate blockages in the coronary arteries and examine heart activity. The salt TlCl can be used, because it acts like NaCl and follows the blood. Gallium-67 accumulates where there is rapid cell growth, such as in tumors and sites of infection. Hence, it is useful in cancer imaging. Usually, the patient receives the injection one day and has a whole body scan 3 or 4 days later because it can take several days for the gallium to build up.

<table>
<thead>
<tr>
<th>Procedure, isotope</th>
<th>Typical activity (mCi), where 1 mCi = 3.7×10^7 Bq</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brain scan</strong></td>
<td></td>
</tr>
<tr>
<td>99m-Tc</td>
<td>7.5</td>
</tr>
<tr>
<td>113m-In</td>
<td>7.5</td>
</tr>
<tr>
<td>11C (PET)</td>
<td>20</td>
</tr>
<tr>
<td>13N (PET)</td>
<td>20</td>
</tr>
<tr>
<td>15O (PET)</td>
<td>50</td>
</tr>
<tr>
<td>18F (PET)</td>
<td>10</td>
</tr>
<tr>
<td><strong>Lung scan</strong></td>
<td></td>
</tr>
<tr>
<td>99m-Tc</td>
<td>2</td>
</tr>
<tr>
<td>133Xe</td>
<td>7.5</td>
</tr>
<tr>
<td><strong>Cardiovascular blood pool</strong></td>
<td></td>
</tr>
<tr>
<td>131I</td>
<td>0.2</td>
</tr>
<tr>
<td>99m-Tc</td>
<td>2</td>
</tr>
<tr>
<td><strong>Cardiovascular arterial flow</strong></td>
<td></td>
</tr>
<tr>
<td>201Tl</td>
<td>3</td>
</tr>
<tr>
<td>24Na</td>
<td>7.5</td>
</tr>
<tr>
<td><strong>Thyroid scan</strong></td>
<td></td>
</tr>
<tr>
<td>131I</td>
<td>0.05</td>
</tr>
<tr>
<td>123I</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Liver scan</strong></td>
<td></td>
</tr>
<tr>
<td>198Au (colloid)</td>
<td>0.1</td>
</tr>
<tr>
<td>99m-Tc (colloid)</td>
<td>2</td>
</tr>
<tr>
<td><strong>Bone scan</strong></td>
<td></td>
</tr>
<tr>
<td>85Sr</td>
<td>0.1</td>
</tr>
<tr>
<td>99m-Tc</td>
<td>10</td>
</tr>
<tr>
<td><strong>Kidney scan</strong></td>
<td></td>
</tr>
<tr>
<td>197Hg</td>
<td>0.1</td>
</tr>
<tr>
<td>99m-Tc</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Note that Table 32.1 lists many diagnostic uses for $^{99m}\text{Tc}$, where “m” stands for a metastable state of the technetium nucleus. Perhaps 80 percent of all radiopharmaceutical procedures employ $^{99m}\text{Tc}$ because of its many advantages. One is that the decay of its metastable state produces a single, easily identified 0.142-MeV $\gamma$ ray. Additionally, the radiation dose to the patient is limited by the short 6.0-h half-life of $^{99m}\text{Tc}$. And, although its half-life is short, it is easily and continuously produced on site. The basic process for production is neutron activation of molybdenum, which quickly $\beta^-$ decays into $^{99m}\text{Tc}$. Technetium-99m can be attached to many compounds to allow the imaging of the skeleton, heart, lungs, kidneys, etc.

Figure 32.5 shows one of the simpler methods of imaging the concentration of nuclear activity, employing a device called an Anger camera or gamma camera. A piece of lead with holes bored through it collimates $\gamma$ rays emerging from the patient, allowing detectors to receive $\gamma$ rays from specific directions only. The computer analysis of detector signals produces an image. One of the disadvantages of this detection method is that there is no depth information (i.e., it provides a two-dimensional view of the tumor as opposed to a three-dimensional view), because radiation from any location under that detector produces a signal.

Imaging techniques much like those in x-ray computed tomography (CT) scans use nuclear activity in patients to form three-dimensional images. Figure 32.6 shows a patient in a circular array of detectors that may be stationary or rotated, with detector output used by a computer to construct a detailed image. This technique is called single-photon-emission computed tomography (SPECT) or sometimes simply SPET. The spatial resolution of this technique is poor, about 1 cm, but the contrast (i.e., the difference in visual properties that makes an object distinguishable from other objects and the background) is good.

Images produced by $\beta^+$ emitters have become important in recent years. When the emitted positron ($\beta^+$) encounters an electron, mutual annihilation occurs, producing two $\gamma$ rays. These $\gamma$ rays have identical 0.511-MeV energies (the energy comes from the destruction of an electron or positron mass) and they move directly away from one another, allowing detectors to determine their point of origin accurately, as shown in Figure 32.7. The system is called positron emission tomography (PET). It requires detectors on opposite sides to simultaneously (i.e., at the same time) detect photons of 0.511-MeV energy and utilizes computer imaging techniques similar to those in SPECT and CT scans. Examples of $\beta^+$-emitting isotopes used in PET are $^{11}\text{C}$, $^{13}\text{N}$, $^{15}\text{O}$, and $^{18}\text{F}$, as seen in Table 32.1. This list includes C, N, and O, and so they have the advantage of being able to function as tags for natural body compounds. Its resolution of 0.5 cm is better than that of SPECT; the accuracy and sensitivity of PET scans make them useful for examining the brain's anatomy and function. The brain's use of oxygen and water can be monitored with $^{15}\text{O}$. PET is used extensively for diagnosing brain disorders. It can note decreased metabolism in certain regions prior to a confirmation of Alzheimer's disease. PET can locate regions in the brain that become active when a person carries out specific activities, such as speaking, closing their eyes, and so on.
A PET system takes advantage of the two identical $\gamma$-ray photons produced by positron-electron annihilation. These $\gamma$ rays are emitted in opposite directions, so that the line along which each pair is emitted is determined. Various events detected by several pairs of detectors are then analyzed by the computer to form an accurate image.

**PhET Explorations: Simplified MRI**

Is it a tumor? Magnetic Resonance Imaging (MRI) can tell. Your head is full of tiny radio transmitters (the nuclear spins of the hydrogen nuclei of your water molecules). In an MRI unit, these little radios can be made to broadcast their positions, giving a detailed picture of the inside of your head.

**PhET Interactive Simulation**

32.2 Biological Effects of Ionizing Radiation

We hear many seemingly contradictory things about the biological effects of ionizing radiation. It can cause cancer, burns, and hair loss, yet it is used to treat and even cure cancer. How do we understand these effects? Once again, there is an underlying simplicity in nature, even in complicated biological organisms. All the effects of ionizing radiation on biological tissue can be understood by knowing that **ionizing radiation affects molecules within cells, particularly DNA molecules**.

Let us take a brief look at molecules within cells and how cells operate. Cells have long, double-helical DNA molecules containing chemical codes called genetic codes that govern the function and processes undertaken by the cell. It is for unraveling the double-helical structure of DNA that James Watson, Francis Crick, and Maurice Wilkins received the Nobel Prize. Damage to DNA consists of breaks in chemical bonds or other changes in the structural features of the DNA chain, leading to changes in the genetic code. In human cells, we can have as many as a million individual instances of damage to DNA per cell per day. It is remarkable that DNA contains codes that check whether the DNA is damaged or can repair itself. It is like an auto check and repair mechanism. This repair ability of DNA is vital for maintaining the integrity of the genetic code and for the normal functioning of the entire organism. It should be constantly active and needs to respond rapidly. The rate of DNA repair depends on various factors such as the cell type and age of the cell. A cell with a damaged ability to repair DNA, which could have been induced by ionizing radiation, can do one of the following:

- The cell can go into an irreversible state of dormancy, known as senescence.
- The cell can commit suicide, known as programmed cell death.
- The cell can go into unregulated cell division leading to tumors and cancers.

Since ionizing radiation damages the DNA, which is critical in cell reproduction, it has its greatest effect on cells that rapidly reproduce, including most types of cancer. Thus, cancer cells are more sensitive to radiation than normal cells and can be killed by it easily. Cancer is characterized by a malfunction of cell reproduction, and can also be caused by ionizing radiation. Without contradiction, ionizing radiation can be both a cure and a cause.

To discuss quantitatively the biological effects of ionizing radiation, we need a radiation dose unit that is directly related to those effects. All effects of radiation are assumed to be directly proportional to the amount of ionization produced in the biological organism. The amount of ionization is in turn proportional to the amount of deposited energy. Therefore, we define a **radiation dose unit** called the rad, as $1/100$ of a joule of ionizing energy deposited per kilogram of tissue, which is

$$1 \text{ rad} = 0.01 \text{ J/kg}.$$  \hspace{1cm} (32.1)

For example, if a 50.0-kg person is exposed to ionizing radiation over her entire body and she absorbs 1.00 J, then her whole-body radiation dose is

$$\frac{1.00 \text{ J}}{50.0 \text{ kg}} = 0.0200 \text{ J/kg} = 2.00 \text{ rad}.$$  \hspace{1cm} (32.2)

If the same 1.00 J of ionizing energy were absorbed in her 2.00-kg forearm alone, then the dose to the forearm would be
1.00 J / (2.00 kg) = 0.500 J/kg = 50.0 rad,

and the unaffected tissue would have a zero rad dose. While calculating radiation doses, you divide the energy absorbed by the mass of affected tissue. You must specify the affected region, such as the whole body or forearm in addition to giving the numerical dose in rads. The SI unit for radiation dose is the **gray (Gy)**, which is defined to be

\[ 1 \text{ Gy} = 1 \text{ J/kg} = 100 \text{ rad}. \] (32.4)

However, the rad is still commonly used. Although the energy per kilogram in 1 rad is small, it has significant effects since the energy causes ionization. The energy needed for a single ionization is a few eV, or less than \(10^{-18} \text{ J}\). Thus, 0.01 J of ionizing energy can create a huge number of ion pairs and have an effect at the cellular level.

The effects of ionizing radiation may be directly proportional to the dose in rads, but they also depend on the type of radiation and the type of tissue. That is, for a given dose in rads, the effects depend on whether the radiation is \(\alpha, \beta, \gamma, \text{x-ray}\), or some other type of ionizing radiation. In the earlier discussion of the range of ionizing radiation, it was noted that energy is deposited in a series of ionizations and not in a single interaction. Each ion pair or ionization requires a certain amount of energy, so that the number of ion pairs is directly proportional to the amount of the deposited ionizing energy. But, if the range of the radiation is small, as it is for \(\alpha\) s, then the ionization and the damage created is more concentrated and harder for the organism to repair, as seen in Figure 32.9. Concentrated damage is more difficult for biological organisms to repair than damage that is spread out, so short-range particles have greater biological effects. The **relative biological effectiveness (RBE) or quality factor (QF)** is given in Table 32.2 for several types of ionizing radiation—the effect of the radiation is directly proportional to the RBE. A dose unit more closely related to effects in biological tissue is called the **roentgen equivalent man** or rem and is defined to be the dose in rads multiplied by the relative biological effectiveness.

\[ \text{rem} = \text{rad} \times \text{RBE} \] (32.5)

Figure 32.9 The image shows ionization created in cells by \(\alpha\) and \(\gamma\) radiation. Because of its shorter range, the ionization and damage created by \(\alpha\) is more concentrated and harder for the organism to repair. Thus, the RBE for \(\alpha\) s is greater than the RBE for \(\gamma\) s, even though they create the same amount of ionization at the same energy.

So, if a person had a whole-body dose of 2.00 rad of \(\gamma\) radiation, the dose in rem would be \((2.00 \text{ rad})(1) = 2.00 \text{ rem whole body}\). If the person had a whole-body dose of 2.00 rad of \(\alpha\) radiation, then the dose in rem would be \((2.00 \text{ rad})(20) = 40.0 \text{ rem whole body}\). The \(\alpha\) s would have 20 times the effect on the person than the \(\gamma\) s for the same deposited energy. The SI equivalent of the rem is the **sievert (Sv)**, defined to be \(\text{Sv} = \text{Gy} \times \text{RBE}\), so that

\[ 1 \text{ Sv} = 1 \text{ Gy} \times \text{RBE} = 100 \text{ rem}. \] (32.6)

The RBEs given in Table 32.2 are approximate, but they yield certain insights. For example, the eyes are more sensitive to radiation, because the cells of the lens do not repair themselves. Neutrons cause more damage than \(\gamma\) rays, although both are neutral and have large ranges, because neutrons often cause secondary radiation when they are captured. Note that the RBEs are 1 for higher-energy \(\beta\) s, \(\gamma\) s, and x-rays, three of the most common types of radiation. For those types of radiation, the numerical values of the dose in rem and rad are identical. For example, 1 rad of \(\gamma\) radiation is also 1 rem. For that reason, rads are still widely quoted rather than rem. Table 32.3 summarizes the units that are used for radiation.

**Miscorception Alert: Activity vs. Dose**

“Activity” refers to the radioactive source while “dose” refers to the amount of energy from the radiation that is deposited in a person or object.

A high level of activity doesn’t mean much if a person is far away from the source. The activity \(R\) of a source depends upon the quantity of material (kg) as well as the half-life. A short half-life will produce many more disintegrations per second. Recall that \(R = \frac{0.693N}{t_{1/2}}\). Also, the activity decreases exponentially, which is seen in the equation \(R = R_0 e^{-\lambda t}\).
Table 32.2 Relative Biological Effectiveness

<table>
<thead>
<tr>
<th>Type and energy of radiation</th>
<th>RBE[^1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-rays</td>
<td>1</td>
</tr>
<tr>
<td>γ rays</td>
<td>1</td>
</tr>
<tr>
<td>β rays greater than 32 keV</td>
<td>1</td>
</tr>
<tr>
<td>β rays less than 32 keV</td>
<td>1.7</td>
</tr>
<tr>
<td>Neutrons, thermal slow (&lt;20 keV)</td>
<td>2–5</td>
</tr>
<tr>
<td>Neutrons, fast (1–10 MeV)</td>
<td>10 (body), 32 (eyes)</td>
</tr>
<tr>
<td>Protons (1–10 MeV)</td>
<td>10 (body), 32 (eyes)</td>
</tr>
<tr>
<td>α rays from radioactive decay</td>
<td>10–20</td>
</tr>
<tr>
<td>Heavy ions from accelerators</td>
<td>10–20</td>
</tr>
</tbody>
</table>

Table 32.3 Units for Radiation

<table>
<thead>
<tr>
<th>Quantity</th>
<th>SI unit name</th>
<th>Definition</th>
<th>Former unit</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>Becquerel (bq)</td>
<td>decay/sec</td>
<td>Curie (Ci)</td>
<td>1 Bq = 2.7×10^{-11} Ci</td>
</tr>
<tr>
<td>Absorbed dose</td>
<td>Gray (Gy)</td>
<td>1 J/kg</td>
<td>rad</td>
<td>Gy = 100 rad</td>
</tr>
<tr>
<td>Dose Equivalent</td>
<td>Sievert (Sv)</td>
<td>1 J/kg × RBE</td>
<td>rem</td>
<td>Sv = 100 rem</td>
</tr>
</tbody>
</table>

The large-scale effects of radiation on humans can be divided into two categories: immediate effects and long-term effects. Table 32.4 gives the immediate effects of whole-body exposures received in less than one day. If the radiation exposure is spread out over more time, greater doses are needed to cause the effects listed. This is due to the body's ability to partially repair the damage. Any dose less than 100 mSv (10 rem) is called a **low dose**, 0.1 Sv to 1 Sv (10 to 100 rem) is called a **moderate dose**, and anything greater than 1 Sv (100 rem) is called a **high dose**. There is no known way to determine after the fact if a person has been exposed to less than 10 mSv.

Table 32.4 Immediate Effects of Radiation (Adults, Whole Body, Single Exposure)

<table>
<thead>
<tr>
<th>Dose in Sv [^2]</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–0.10</td>
<td>No observable effect.</td>
</tr>
<tr>
<td>0.1 – 1</td>
<td>Slight to moderate decrease in white blood cell counts.</td>
</tr>
<tr>
<td>0.5</td>
<td>Temporary sterility; 0.35 for women, 0.50 for men.</td>
</tr>
<tr>
<td>1 – 2</td>
<td>Significant reduction in blood cell counts, brief nausea and vomiting. Rarely fatal.</td>
</tr>
<tr>
<td>2 – 5</td>
<td>Nausea, vomiting, hair loss, severe blood damage, hemorrhage, fatalities.</td>
</tr>
<tr>
<td>4.5</td>
<td>LD50/32. Lethal to 50% of the population within 32 days after exposure if not treated.</td>
</tr>
<tr>
<td>5 – 20</td>
<td>Worst effects due to malfunction of small intestine and blood systems. Limited survival.</td>
</tr>
<tr>
<td>&gt;20</td>
<td>Fatal within hours due to collapse of central nervous system.</td>
</tr>
</tbody>
</table>

Immediate effects are explained by the effects of radiation on cells and the sensitivity of rapidly reproducing cells to radiation. The first clue that a person has been exposed to radiation is a change in blood count, which is not surprising since blood cells are the most rapidly reproducing cells in the body. At higher doses, nausea and hair loss are observed, which may be due to interference with cell reproduction. Cells in the lining of the digestive system also rapidly reproduce, and their destruction causes nausea. When the growth of hair cells slows, the hair follicles become thin and break off. High doses cause significant cell death in all systems, but the lowest doses that cause fatalities do so by weakening the immune system through the loss of white blood cells.

The two known long-term effects of radiation are cancer and genetic defects. Both are directly attributable to the interference of radiation with cell reproduction. For high doses of radiation, the risk of cancer is reasonably well known from studies of exposed groups. Hiroshima and Nagasaki survivors and a smaller number of people exposed by their occupation, such as radium dial painters, have been fully documented. Chernobyl victims will be studied for many decades, with some data already available. For example, a significant increase in childhood thyroid cancer has been observed. The risk of a radiation-induced cancer for low and moderate doses is generally **assumed** to be proportional to the risk known for high doses. Under this assumption, any dose of radiation, no matter how small, involves a risk to human health. This is called the **linear hypothesis** and it may be prudent, but it is controversial. There is some evidence that, unlike the immediate effects of radiation, the long-term effects are cumulative and there is little self-repair. This is analogous to the risk of skin cancer from UV exposure, which is known to be cumulative.

There is a latency period for the onset of radiation-induced cancer of about 2 years for leukemia and 15 years for most other forms. The person is at risk for at least 30 years after the latency period. Omitting many details, the overall risk of a radiation-induced cancer death per year per rem of exposure is about 10 in a million, which can be written as $10 / 10^6$ rem · y.

[^1]: Values approximate, difficult to determine.
[^2]: Multiply by 100 to obtain dose in rem.
If a person receives a dose of 1 rem, his risk each year of dying from radiation-induced cancer is 10 in a million and that risk continues for about 30 years. The lifetime risk is thus 300 in a million, or 0.03 percent. Since about 20 percent of all worldwide deaths are from cancer, the increase due to a 1 rem exposure is impossible to detect demographically. But 100 rem (1 Sv), which was the dose received by the average Hiroshima and Nagasaki survivor, causes a 3 percent risk, which can be observed in the presence of a 20 percent normal or natural incidence rate.

The incidence of genetic defects induced by radiation is about one-third that of cancer deaths, but is much more poorly known. The lifetime risk of a genetic defect due to a 1 rem exposure is about 100 in a million or \(3.3 \times 10^{-6}\) rem \(\cdot\) y, but the normal incidence is 60,000 in a million. Evidence of such a small increase, tragic as it is, is nearly impossible to obtain. For example, there is no evidence of increased genetic defects among the offspring of Hiroshima and Nagasaki survivors. Animal studies do not seem to correlate well with effects on humans and are not very helpful. For both cancer and genetic defects, the approach to safety has been to use the linear hypothesis, which is likely to be an overestimate of the risks of low doses. Certain researchers even claim that low doses are beneficial. Hormesis is a term used to describe generally favorable biological responses to low exposures of toxins or radiation. Such low levels may help certain repair mechanisms to develop or enable cells to adapt to the effects of the low exposures. Positive effects may occur at low doses that could be a problem at high doses.

Even the linear hypothesis estimates of the risks are relatively small, and the average person is not exposed to large amounts of radiation. Table 32.5 lists average annual background radiation doses from natural and artificial sources for Australia, the United States, Germany, and world-wide averages. Cosmic rays are partially shielded by the atmosphere, and the dose depends upon altitude and latitude, but the average is about 0.40 mSv/y. A good example of the variation of cosmic radiation dose with altitude comes from the airline industry. Monitored personnel show an average of 2 mSv/y. A 12-hour flight might give you an exposure of 0.02 to 0.03 mSv.

Doses from the Earth itself are mainly due to the isotopes of uranium, thorium, and potassium, and vary greatly by location. Some places have great natural concentrations of uranium and thorium, yielding doses ten times as high as the average value. Internal doses come from foods and liquids that we ingest. Fertilizers containing phosphates have potassium and uranium. So we are all a little radioactive. Carbon-14 has about 66 Bq/kg radioactivity whereas fertilizers may have more than 3000 Bq/kg radioactivity. Medical and dental diagnostic exposures are mostly from x-rays. It should be noted that x-ray doses tend to be localized and are becoming much smaller with improved techniques. Table 32.6 shows typical doses received during various diagnostic x-ray examinations. Note the large dose from a CT scan. While CT scans only account for less than 20 percent of the x-ray procedures done today, they account for about 50 percent of the annual dose received.

Radon is usually more pronounced underground and in buildings with low air exchange with the outside world. Almost all soil contains some \(^{226}\text{Ra}\) and \(^{222}\text{Rn}\), but radon is lower in mainly sedimentary soils and higher in granite soils. Thus, the exposure to the public can vary greatly, even within short distances. Radon can diffuse from the soil into homes, especially basements. The estimated exposure for \(^{222}\text{Rn}\) is controversial. Recent studies indicate there is more radon in homes than had been realized, and it is speculated that radon may be responsible for 20 percent of lung cancers, being particularly hazardous to those who also smoke. Many countries have introduced limits on allowable radon concentrations in indoor air, often requiring the measurement of radon concentrations in a house prior to its sale. Ironically, it could be argued that the higher levels of radon exposure and their geographic variability, taken with the lack of demographic evidence of any effects, means that low-level radiation is less dangerous than previously thought.

**Radiation Protection**

Laws regulate radiation doses to which people can be exposed. The greatest occupational whole-body dose that is allowed depends upon the country and is about 20 to 50 mSv/y and is rarely reached by medical and nuclear power workers. Higher doses are allowed for the hands. Much lower doses are permitted for the reproductive organs and the fetuses of pregnant women. Inadvertent doses to the public are limited to 0.05 mSv/y (5 mrem/y). This has been exceeded in the United States only at the time of the Three Mile Island (TMI) accident in 1979.

Chernobyl is another story. Extensive monitoring with a variety of radiation detectors is performed to assure radiation safety. Increased ventilation in uranium mines has lowered the dose there to about 1 mSv.

### Table 32.5 Background Radiation Sources and Average Doses

<table>
<thead>
<tr>
<th>Source</th>
<th>Dose (mSv/y)</th>
<th>Source</th>
<th>Dose (mSv/y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural Radiation - external</td>
<td></td>
<td>Natural Radiation - internal</td>
<td></td>
</tr>
<tr>
<td>Cosmic Rays</td>
<td>0.30</td>
<td>(^{40}\text{K}, :^{14}\text{C}, :^{226}\text{Ra}</td>
<td>0.24</td>
</tr>
<tr>
<td>Soil, building materials</td>
<td>0.40</td>
<td>Medical &amp; Dental</td>
<td>0.80</td>
</tr>
<tr>
<td>Radon gas</td>
<td>0.90</td>
<td></td>
<td>0.90</td>
</tr>
<tr>
<td>Total</td>
<td>2.6</td>
<td></td>
<td>3.5</td>
</tr>
</tbody>
</table>

To physically limit radiation doses, we use **shielding**, increase the **distance** from a source, and limit the **time of exposure**.

**Figure 32.10** illustrates how these are used to protect both the patient and the dental technician when an x-ray is taken. Shielding absorbs radiation and can be provided by any material, including sufficient air. The greater the distance from the source, the more the radiation spreads out. The less

---

3. Multiply by 100 to obtain dose in mrem/y.
time a person is exposed to a given source, the smaller is the dose received by the person. Doses from most medical diagnostics have decreased in recent years due to faster films that require less exposure time.

![Figure 32.10](image)

Figure 32.10 A lead apron is placed over the dental patient and shielding surrounds the x-ray tube to limit exposure to tissue other than the tissue that is being imaged. Fast films limit the time needed to obtain images, reducing exposure to the imaged tissue. The technician stands a few meters away behind a lead-lined door with a lead glass window, reducing her occupational exposure.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Effective dose (mSv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest</td>
<td>0.02</td>
</tr>
<tr>
<td>Dental</td>
<td>0.01</td>
</tr>
<tr>
<td>Skull</td>
<td>0.07</td>
</tr>
<tr>
<td>Leg</td>
<td>0.02</td>
</tr>
<tr>
<td>Mammogram</td>
<td>0.40</td>
</tr>
<tr>
<td>Barium enema</td>
<td>7.0</td>
</tr>
<tr>
<td>Upper GI</td>
<td>3.0</td>
</tr>
<tr>
<td>CT head</td>
<td>2.0</td>
</tr>
<tr>
<td>CT abdomen</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Table 32.6 Typical Doses Received During Diagnostic X-ray Exams

**Problem-Solving Strategy**  
You need to follow certain steps for dose calculations, which are

**Step 1.** Examine the situation to determine that a person is exposed to ionizing radiation.

**Step 2.** Identify exactly what needs to be determined in the problem (identify the unknowns). The most straightforward problems ask for a dose calculation.

**Step 3.** Make a list of what is given or can be inferred from the problem as stated (identify the knowns). Look for information on the type of radiation, the energy per event, the activity, and the mass of tissue affected.

**Step 4.** For dose calculations, you need to determine the energy deposited. This may take one or more steps, depending on the given information.

**Step 5.** Divide the deposited energy by the mass of the affected tissue. Use units of joules for energy and kilograms for mass. If a dose in Sv is involved, use the definition that $1\text{ Sv} = 1\text{ J/kg}$.

**Step 6.** If a dose in mSv is involved, determine the RBE (QF) of the radiation. Recall that $1\text{ mSv} = 1\text{ mGy} \times \text{RBE}$ (or $1\text{ rem} = 1\text{ rad} \times \text{RBE}$).

**Step 7.** Check the answer to see if it is reasonable: Does it make sense? The dose should be consistent with the numbers given in the text for diagnostic, occupational, and therapeutic exposures.

**Example 32.1 Dose from Inhaled Plutonium**

Calculate the dose in rem/y for the lungs of a weapons plant employee who inhales and retains an activity of $1.00 \mu\text{Ci}$ of $^{239}\text{Pu}$ in an accident. The mass of affected lung tissue is $2.00$ kg, the plutonium decays by emission of a $5.23\text{-MeV} \alpha$ particle, and you may assume the higher value of the RBE for $\alpha$ s from Table 32.2.

**Strategy**

Dose in rem is defined by $1\text{ rad} = 0.01\text{ J/kg}$ and $1\text{ rem} = \text{rad} \times \text{RBE}$. The energy deposited is divided by the mass of tissue affected and then multiplied by the RBE. The latter two quantities are given, and so the main task in this example will be to find the energy deposited in one year.
Since the activity of the source is given, we can calculate the number of decays, multiply by the energy per decay, and convert MeV to joules to get the total energy.

**Solution**

The activity \( R = 1.00 \mu\text{Ci} = 3.70 \times 10^4 \text{ Bq} = 3.70 \times 10^4 \text{ decays/s} \). So, the number of decays per year is obtained by multiplying by the number of seconds in a year:

\[
(3.70 \times 10^4 \text{ decays/s})(3.16 \times 10^7 \text{ s}) = 1.17 \times 10^{12} \text{ decays}.
\]

Thus, the ionizing energy deposited per year is

\[
E = (1.17 \times 10^{12} \text{ decays})(5.23 \text{ MeV/decay})(1.60 \times 10^{-13} \text{ J/MeV}) = 0.978 \text{ J}.
\]

Dividing by the mass of the affected tissue gives

\[
\frac{E}{\text{mass}} = \frac{0.978 \text{ J}}{2.00 \text{ kg}} = 0.489 \text{ J/kg}.
\]

One Gray is 1.00 J/kg, and so the dose in Gy is

\[
\text{dose in Gy} = \frac{0.489 \text{ J/kg}}{1.00 \text{ (J/kg)/Gy}} = 0.489 \text{ Gy}.
\]

Now, the dose in Sv is

\[
\text{dose in Sv} = \text{Gy} \times \text{RBE} = (0.489 \text{ Gy})(20) = 9.8 \text{ Sv}.
\]

**Discussion**

First note that the dose is given to two digits, because the RBE is (at best) known only to two digits. By any standard, this yearly radiation dose is high and will have a devastating effect on the health of the worker. Worse yet, plutonium has a long radioactive half-life and is not readily eliminated by the body, and so it will remain in the lungs. Being an \( \alpha \) emitter makes the effects 10 to 20 times worse than the same ionization produced by \( \beta \) s, \( \gamma \) rays, or \( x \)-rays. An activity of 1.00 \( \mu\text{Ci} \) is created by only 16 \( \mu\text{g} \) of \( ^{239}\text{Pu} \) (left as an end-of-chapter problem to verify), partly justifying claims that plutonium is the most toxic substance known. Its actual hazard depends on how likely it is to be spread out among a large population and then ingested. The Chernobyl disaster’s deadly legacy, for example, has nothing to do with the plutonium it put into the environment.

**Risk versus Benefit**

Medical doses of radiation are also limited. Diagnostic doses are generally low and have further lowered with improved techniques and faster films. With the possible exception of routine dental x-rays, radiation is used diagnostically only when needed so that the low risk is justified by the benefit of the diagnosis. Chest x-rays give the lowest doses—about 0.1 mSv to the tissue affected, with less than 5 percent scattering into tissues that are not directly imaged. Other x-ray procedures range upward to about 10 mSv in a CT scan, and about 5 mSv (0.5 rem) per dental x-ray, again both only affecting the tissue imaged. Medical images with radiopharmaceuticals give doses ranging from 1 to 5 mSv, usually localized. One exception is the thyroid scan using \( ^{131}\text{I} \). Because of its relatively long half-life, it exposes the thyroid to about 0.75 Sv. The isotope \( ^{123}\text{I} \) is more difficult to produce, but its short half-life limits thyroid exposure to about 15 mSv.

**PhET Explorations: Alpha Decay**

Watch alpha particles escape from a polonium nucleus, causing radioactive alpha decay. See how random decay times relate to the half life.

**Figure 32.11 Alpha Decay**

**32.3 Therapeutic Uses of Ionizing Radiation**

Therapeutic applications of ionizing radiation, called radiation therapy or radiotherapy, have existed since the discovery of x-rays and nuclear radioactivity. Today, radiotherapy is used almost exclusively for cancer therapy, where it saves thousands of lives and improves the quality of life and longevity of many it cannot save. Radiotherapy may be used alone or in combination with surgery and chemotherapy (drug treatment) depending on the type of cancer and the response of the patient. A careful examination of all available data has established that radiotherapy’s beneficial effects far outweigh its long-term risks.
Medical Application

The earliest uses of ionizing radiation on humans were mostly harmful, with many at the level of snake oil as seen in Figure 32.12. Radium-doped cosmetics that glowed in the dark were used around the time of World War I. As recently as the 1950s, radon mine tours were promoted as healthful and rejuvenating—those who toured were exposed but gained no benefits. Radium salts were sold as health elixirs for many years. The gruesome death of a wealthy industrialist, who became psychologically addicted to the brew, alerted the unsuspecting to the dangers of radium salt elixirs. Most abuses finally ended after the legislation in the 1950s.

Radiotherapy is effective against cancer because cancer cells reproduce rapidly and, consequently, are more sensitive to radiation. The central problem in radiotherapy is to make the dose for cancer cells as high as possible while limiting the dose for normal cells. The ratio of abnormal cells killed to normal cells killed is called the therapeutic ratio, and all radiotherapy techniques are designed to enhance this ratio. Radiation can be concentrated in cancerous tissue by a number of techniques. One of the most prevalent techniques for well-defined tumors is a geometric technique shown in Figure 32.13. A narrow beam of radiation is passed through the patient from a variety of directions with a common crossing point in the tumor. This concentrates the dose in the tumor while spreading it out over a large volume of normal tissue. The external radiation can be x-rays, $^{60}$Co $\gamma$ rays, or ionizing-particle beams produced by accelerators. Accelerator-produced beams of neutrons, $\pi$-mesons, and heavy ions such as nitrogen nuclei have been employed, and these can be quite effective. These particles have larger QFs or RBEs and sometimes can be better localized, producing a greater therapeutic ratio. But accelerator radiotherapy is much more expensive and less frequently employed than other forms.

Figure 32.12 The properties of radiation were once touted for far more than its modern use in cancer therapy. Until 1932, radium was advertised for a variety of uses, often with tragic results. (credit: Struthious Bandersnatch.)

Figure 32.13 The $^{60}$Co source of $\gamma$-radiation is rotated around the patient so that the common crossing point is in the tumor, concentrating the dose there. This geometric technique works for well-defined tumors.
Another form of radiotherapy uses chemically inert radioactive implants. One use is for prostate cancer. Radioactive seeds (about 40 to 100 and the size of a grain of rice) are placed in the prostate region. The isotopes used are usually $^{135}$I (6-month half life) or $^{103}$Pd (3-month half life). Alpha emitters have the dual advantages of a large QF and a small range for better localization.

Radiopharmaceuticals are used for cancer therapy when they can be localized well enough to produce a favorable therapeutic ratio. Thyroid cancer is commonly treated utilizing radioactive iodine. Thyroid cells concentrate iodine, and cancerous thyroid cells are more aggressive in doing this. An ingenious use of radiopharmaceuticals in cancer therapy tags antibodies with radioisotopes. Antibodies produced by a patient to combat his cancer are extracted, cultured, loaded with a radioisotope, and then returned to the patient. The antibodies are concentrated almost entirely in the tissue they developed to fight, thus localizing the radiation in abnormal tissue. The therapeutic ratio can be quite high for short-range radiation. There is, however, a significant dose for organs that eliminate radiopharmaceuticals from the body, such as the liver, kidneys, and bladder. As with most radiotherapy, the technique is limited by the tolerable amount of damage to the normal tissue.

Table 32.7 lists typical therapeutic doses of radiation used against certain cancers. The doses are large, but not fatal because they are localized and spread out in time. Protocols for treatment vary with the type of cancer and the condition and response of the patient. Three to five 200-rem treatments per week for a period of several weeks is typical. Time between treatments allows the body to repair normal tissue. This effect occurs because damage is concentrated in the abnormal tissue, and the abnormal tissue is more sensitive to radiation. Damage to normal tissue limits the doses. You will note that the greatest doses are given to any tissue that is not rapidly reproducing, such as in the adult brain. Lung cancer, on the other end of the scale, cannot ordinarily be cured with radiation because of the sensitivity of lung tissue and blood to radiation. But radiotherapy for lung cancer does alleviate symptoms and prolong life and is therefore justified in some cases.

<table>
<thead>
<tr>
<th>Type of Cancer</th>
<th>Typical dose (Sv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>10–20</td>
</tr>
<tr>
<td>Hodgkin’s disease</td>
<td>40–45</td>
</tr>
<tr>
<td>Skin</td>
<td>40–50</td>
</tr>
<tr>
<td>Ovarian</td>
<td>50–75</td>
</tr>
<tr>
<td>Breast</td>
<td>50–80+</td>
</tr>
<tr>
<td>Brain</td>
<td>80+</td>
</tr>
<tr>
<td>Neck</td>
<td>80+</td>
</tr>
<tr>
<td>Bone</td>
<td>80+</td>
</tr>
<tr>
<td>Soft tissue</td>
<td>80+</td>
</tr>
<tr>
<td>Thyroid</td>
<td>80+</td>
</tr>
</tbody>
</table>

Finally, it is interesting to note that chemotherapy employs drugs that interfere with cell division and is, thus, also effective against cancer. It also has almost the same side effects, such as nausea and hair loss, and risks, such as the inducement of another cancer.

### 32.4 Food Irradiation

Ionizing radiation is widely used to sterilize medical supplies, such as bandages, and consumer products, such as tampons. Worldwide, it is also used to irradiate food, an application that promises to grow in the future. **Food irradiation** is the treatment of food with ionizing radiation. It is used to reduce pest infestation and to delay spoilage and prevent illness caused by microorganisms. Food irradiation is controversial. Proponents see it as superior to pasteurization, preservatives, and insecticides, supplanting dangerous chemicals with a more effective process. Opponents see its safety as unproven, perhaps leaving worse toxic residues as well as presenting an environmental hazard at treatment sites. In developing countries, food irradiation might increase crop production by 25.0% or more, and reduce food spoilage by a similar amount. It is used chiefly to treat spices and some fruits, and in some countries, red meat, poultry, and vegetables. Over 40 countries have approved food irradiation at some level.

Food irradiation exposes food to large doses of $\gamma$ rays, x-rays, or electrons. These photons and electrons induce no nuclear reactions and thus create no residual radioactivity. (Some forms of ionizing radiation, such as neutron irradiation, create residual radioactivity. These are not used for food irradiation.) The $\gamma$ source is usually $^{60}$Co or $^{137}$Cs, the latter isotope being a major by-product of nuclear power. Cobalt-60 $\gamma$ rays average 1.25 MeV, while those of $^{137}$Cs are 0.67 MeV and are less penetrating. X-rays used for food irradiation are created with voltages of up to 5 million volts and, thus, have photon energies up to 5 MeV. Electrons used for food irradiation are accelerated to energies up to 10 MeV. The higher the energy per particle, the more penetrating the radiation is and the more ionization it can create. Figure 32.14 shows a typical $\gamma$-irradiation plant.
A food irradiation plant has a conveyor system to pass items through an intense radiation field behind thick shielding walls. The $\gamma$ source is lowered into a deep pool of water for safe storage when not in use. Exposure times of up to an hour expose food to doses up to $10^4$ Gy.

Owing to the fact that food irradiation seeks to destroy organisms such as insects and bacteria, much larger doses than those fatal to humans must be applied. Generally, the simpler the organism, the more radiation it can tolerate. (Cancer cells are a partial exception, because they are rapidly reproducing and, thus, more sensitive.) Current licensing allows up to 1000 Gy to be applied to fresh fruits and vegetables, called a low dose in food irradiation. Such a dose is enough to prevent or reduce the growth of many microorganisms, but about 10,000 Gy is needed to kill salmonella, and even more is needed to kill fungi. Doses greater than 10,000 Gy are considered to be high doses in food irradiation and product sterilization.

The effectiveness of food irradiation varies with the type of food. Spices and many fruits and vegetables have dramatically longer shelf lives. These also show no degradation in taste and no loss of food value or vitamins. If not for the mandatory labeling, such foods subjected to low-level irradiation (up to 1000 Gy) could not be distinguished from untreated foods in quality. However, some foods actually spoil faster after irradiation, particularly those with high water content like lettuce and peaches. Others, such as milk, are given a noticeably unpleasant taste. High-level irradiation produces significant and chemically measurable changes in foods. It produces about a 15% loss of nutrients and a 25% loss of vitamins, as well as some change in taste. Such losses are similar to those that occur in ordinary freezing and cooking.

How does food irradiation work? Ionization produces a random assortment of broken molecules and ions, some with unstable oxygen- or hydrogen-containing molecules known as free radicals. These undergo rapid chemical reactions, producing perhaps four or five thousand different compounds called radiolytic products, some of which make cell function impossible by breaking cell membranes, fracturing DNA, and so on. How safe is the food afterward? Critics argue that the radiolytic products present a lasting hazard, perhaps being carcinogenic. However, the safety of irradiated food is not known precisely. We do know that low-level food irradiation produces no compounds in amounts that can be measured chemically. This is not surprising, since trace amounts of several thousand compounds may be created. We also know that there have been no observable negative short-term effects on consumers. Long-term effects may show up if large number of people consume large quantities of irradiated food, but no effects have appeared due to the small amounts of irradiated food that are consumed regularly. The case for safety is supported by testing of animal diets that were irradiated; no transmitted genetic effects have been observed. Food irradiation (at least up to a million rad) has been endorsed by the World Health Organization and the UN Food and Agricultural Organization. Finally, the hazard to consumers, if it exists, must be weighed against the benefits in food production and preservation. It must also be weighed against the very real hazards of existing insecticides and food preservatives.

### 32.5 Fusion

While basking in the warmth of the summer sun, a student reads of the latest breakthrough in achieving sustained thermonuclear power and vaguely recalls hearing about the cold fusion controversy. The three are connected. The Sun’s energy is produced by nuclear fusion (see Figure 32.15).

Nuclear fusion is a reaction in which two nuclei are combined, or fused, to form a larger nucleus. We know that all nuclei have less mass than the sum of the masses of the protons and neutrons that form them. The missing mass times $c^2$ equals the binding energy of the nucleus—the greater the binding energy, the greater the missing mass. We also know that $BE/A$, the binding energy per nucleon, is greater for medium-mass nuclei and has a maximum at Fe (iron). This means that if two low-mass nuclei can be fused together to form a larger nucleus, energy can be released. The
larger nucleus has a greater binding energy and less mass per nucleon than the two that combined. Thus mass is destroyed in the fusion reaction, and energy is released (see Figure 32.16). On average, fusion of low-mass nuclei releases energy, but the details depend on the actual nuclides involved.

The major obstruction to fusion is the Coulomb repulsion between nuclei. Since the attractive nuclear force that can fuse nuclei together is short ranged, the repulsion of like positive charges must be overcome to get nuclei close enough to induce fusion. Figure 32.17 shows an approximate graph of the potential energy between two nuclei as a function of the distance between their centers. The graph is analogous to a hill with a well in its center. A ball rolled from the right must have enough kinetic energy to get over the hump before it falls into the deeper well with a net gain in energy. So it is with fusion. If the nuclei are given enough kinetic energy to overcome the electric potential energy due to repulsion, then they can combine, release energy, and fall into a deep well. One way to accomplish this is to heat fusion fuel to high temperatures so that the kinetic energy of thermal motion is sufficient to get the nuclei together.

You might think that, in the core of our Sun, nuclei are coming into contact and fusing. However, in fact, temperatures on the order of $10^8$ K are needed to actually get the nuclei in contact, exceeding the core temperature of the Sun. Quantum mechanical tunneling is what makes fusion in the Sun possible, and tunneling is an important process in most other practical applications of fusion, too. Since the probability of tunneling is extremely sensitive to barrier height and width, increasing the temperature greatly increases the rate of fusion. The closer reactants get to one another, the more likely they are to fuse (see Figure 32.18). Thus most fusion in the Sun and other stars takes place at their centers, where temperatures are highest. Moreover, high temperature is needed for thermonuclear power to be a practical source of energy.
The Sun produces energy by fusing protons or hydrogen nuclei $^1\text{H}$ (by far the Sun’s most abundant nuclide) into helium nuclei $^4\text{He}$.

The principal sequence of fusion reactions forms what is called the **proton-proton cycle**:

\[
\begin{align*}
^1\text{H} + ^1\text{H} &\rightarrow ^2\text{H} + e^+ + \nu_e \quad (0.42 \text{ MeV}) \\
^1\text{H} + ^2\text{H} &\rightarrow ^3\text{He} + \gamma \quad (5.49 \text{ MeV}) \\
^3\text{He} + ^3\text{He} &\rightarrow ^4\text{He} + ^1\text{H} + ^1\text{H} \quad (12.86 \text{ MeV})
\end{align*}
\]

where $e^+$ stands for a positron and $\nu_e$ is an electron neutrino. (The energy in parentheses is released by the reaction.) Note that the first two reactions must occur twice for the third to be possible, so that the cycle consumes six protons ($^1\text{H}$) but gives back two. Furthermore, the two positrons produced will find two electrons and annihilate to form four more $\gamma$ rays, for a total of six. The overall effect of the cycle is thus

\[
2e^- + 4^1\text{H} \rightarrow ^4\text{He} + 2\nu_e + 6\gamma \quad (26.7 \text{ MeV})
\]

where the 26.7 MeV includes the annihilation energy of the positrons and electrons and is distributed among all the reaction products. The solar interior is dense, and the reactions occur deep in the Sun where temperatures are highest. It takes about 32,000 years for the energy to diffuse to the surface and radiate away. However, the neutrinos escape the Sun in less than two seconds, carrying their energy with them, because they interact so weakly that the Sun is transparent to them. Negative feedback in the Sun acts as a thermostat to regulate the overall energy output. For instance, if the interior of the Sun becomes hotter than normal, the reaction rate increases, producing energy that expands the interior. This cools it and lowers the reaction rate. Conversely, if the interior becomes too cool, it contracts, increasing the temperature and reaction rate (see Figure 32.19). Stars like the Sun are stable for billions of years, until a significant fraction of their hydrogen has been depleted. What happens then is discussed in *Introduction to Frontiers of Physics*.

Theories of the proton-proton cycle (and other energy-producing cycles in stars) were pioneered by the German-born, American physicist Hans Bethe (1906–2005), starting in 1938. He was awarded the 1967 Nobel Prize in physics for this work, and he has made many other contributions to physics and society. Neutrinos produced in these cycles escape so readily that they provide us an excellent means to test these theories and study stellar interiors. Detectors have been constructed and operated for more than four decades now to measure solar neutrinos (see Figure 32.20). Although solar neutrinos are detected and neutrinos were observed from Supernova 1987A (Figure 32.21), too few solar neutrinos were observed to be consistent with predictions of solar energy production. After many years, this solar neutrino problem was resolved with a blend of theory and experiment that showed that the neutrino does indeed have mass. It was also found that there are three types of neutrinos, each associated with a different type of nuclear decay.
Figure 32.20 This array of photomultiplier tubes is part of the large solar neutrino detector at the Fermi National Accelerator Laboratory in Illinois. In these experiments, the neutrinos interact with heavy water and produce flashes of light, which are detected by the photomultiplier tubes. In spite of its size and the huge flux of neutrinos that strike it, very few are detected each day since they interact so weakly. This, of course, is the same reason they escape the Sun so readily. (credit: Fred Ulrich)

Figure 32.21 Supernovas are the source of elements heavier than iron. Energy released powers nucleosynthesis. Spectroscopic analysis of the ring of material ejected by Supernova 1987A observable in the southern hemisphere, shows evidence of heavy elements. The study of this supernova also provided indications that neutrinos might have mass. (credit: NASA, ESA, and P. Challis)

The proton-proton cycle is not a practical source of energy on Earth, in spite of the great abundance of hydrogen \((^1\text{H})\). The reaction

\[
^1\text{H} + ^1\text{H} \rightarrow ^2\text{H} + e^+ + \nu_e
\]

has a very low probability of occurring. (This is why our Sun will last for about ten billion years.) However, a number of other fusion reactions are easier to induce. Among them are:

\[
\begin{align*}
^2\text{H} + ^2\text{H} & \rightarrow ^3\text{H} + ^1\text{H} \quad (4.03 \text{ MeV}) \\
^2\text{H} + ^3\text{H} & \rightarrow ^3\text{He} + n \quad (3.27 \text{ MeV}) \\
^2\text{H} + ^3\text{He} & \rightarrow ^4\text{He} + n \quad (17.59 \text{ MeV}) \\
^2\text{H} + ^2\text{H} & \rightarrow ^4\text{He} + \gamma \quad (23.85 \text{ MeV}).
\end{align*}
\]

(32.17)  
(32.18)  
(32.19)  
(32.20)

Deuterium \((^2\text{H})\) is about 0.015\% of natural hydrogen, so there is an immense amount of it in sea water alone. In addition to an abundance of deuterium fuel, these fusion reactions produce large energies per reaction (in parentheses), but they do not produce much radioactive waste. Tritium \((^3\text{H})\) is radioactive, but it is consumed as a fuel (the reaction \(^2\text{H} + ^3\text{H} \rightarrow ^4\text{He} + n\)), and the neutrons and \(\gamma\)s can be shielded. The neutrons produced can also be used to create more energy and fuel in reactions like

\[
n + ^1\text{H} \rightarrow ^2\text{H} + \gamma \quad (20.68 \text{ MeV})
\]

(32.21)  

and

\[
n + ^1\text{H} \rightarrow ^2\text{H} + \gamma \quad (2.22 \text{ MeV}).
\]

(32.22)

Note that these last two reactions, and \(^2\text{H} + ^2\text{H} \rightarrow ^4\text{He} + \gamma\), put most of their energy output into the \(\gamma\) ray, and such energy is difficult to utilize.

The three keys to practical fusion energy generation are to achieve the temperatures necessary to make the reactions likely, to raise the density of the fuel, and to confine it long enough to produce large amounts of energy. These three factors—temperature, density, and time—complement one another, and so a deficiency in one can be compensated for by the others. Ignition is defined to occur when the reactions produce enough energy to be self-sustaining after external energy input is cut off. This goal, which must be reached before commercial plants can be a reality, has not been achieved. Another milestone, called break-even, occurs when the fusion power produced equals the heating power input. Break-even has nearly been reached and gives hope that ignition and commercial plants may become a reality in a few decades.

Two techniques have shown considerable promise. The first of these is called magnetic confinement and uses the property that charged particles have difficulty crossing magnetic field lines. The tokamak, shown in Figure 32.22, has shown particular promise. The tokamak’s toroidal coil confines charged particles into a circular path with a helical twist due to the circulating ions themselves. In 1995, the Tokamak Fusion Test Reactor at Princeton in the US achieved world-record plasma temperatures as high as 500 million degrees Celsius. This facility operated between 1982 and 1997. A joint international effort is underway in France to build a tokamak-type reactor that will be the stepping stone to commercial power. ITER, as it is called, will be a full-scale device that aims to demonstrate the feasibility of fusion energy. It will generate 500 MW of power for extended periods of
time and will achieve break-even conditions. It will study plasmas in conditions similar to those expected in a fusion power plant. Completion is scheduled for 2018.

![Figure 32.22](a) Artist’s rendition of ITER, a tokamak-type fusion reactor being built in southern France. It is hoped that this gigantic machine will reach the break-even point. Completion is scheduled for 2018. (credit: Stephan Mosel, Flickr)

The second promising technique aims multiple lasers at tiny fuel pellets filled with a mixture of deuterium and tritium. Huge power input heats the fuel, evaporating the confining pellet and crushing the fuel to high density with the expanding hot plasma produced. This technique is called **inertial confinement**, because the fuel's inertia prevents it from escaping before significant fusion can take place. Higher densities have been reached than with tokamaks, but with smaller confinement times. In 2009, the Lawrence Livermore Laboratory (CA) completed a laser fusion device with 192 ultraviolet laser beams that are focused upon a D-T pellet (see Figure 32.23).

![Figure 32.23](a) National Ignition Facility (CA). This image shows a laser bay where 192 laser beams will focus onto a small D-T target, producing fusion. (credit: Lawrence Livermore National Laboratory, Lawrence Livermore National Security, LLC, and the Department of Energy)

**Example 32.2 Calculating Energy and Power from Fusion**

(a) Calculate the energy released by the fusion of a 1.00-kg mixture of deuterium and tritium, which produces helium. There are equal numbers of deuterium and tritium nuclei in the mixture.

(b) If this takes place continuously over a period of a year, what is the average power output?

**Strategy**

According to \( ^2 \text{H} + ^3 \text{H} \rightarrow ^4 \text{He} + n \), the energy per reaction is 17.59 MeV. To find the total energy released, we must find the number of deuterium and tritium atoms in a kilogram. Deuterium has an atomic mass of about 2 and tritium has an atomic mass of about 3, for a total of about 5 g per mole of reactants or about 200 mol in 1.00 kg. To get a more precise figure, we will use the atomic masses from Appendix A. The power output is best expressed in watts, and so the energy output needs to be calculated in joules and then divided by the number of seconds in a year.

**Solution for (a)**

The atomic mass of deuterium (\( ^2 \text{H} \)) is 2.014102 u, while that of tritium (\( ^3 \text{H} \)) is 3.016049 u, for a total of 5.032151 u per reaction. So a mole of reactants has a mass of 5.03 g, and in 1.00 kg there are \( (1000 \text{ g}) / (5.03 \text{ g/mol}) = 198.8 \text{ mol of reactants} \). The number of reactions that take place is therefore

\[
(198.8 \text{ mol})(6.02\times10^{23} \text{ mol}^{-1}) = 1.20\times10^{26} \text{ reactions.}
\]  

(32.23)

The total energy output is the number of reactions times the energy per reaction:

\[
E = (1.20\times10^{26} \text{ reactions})(17.59 \text{ MeV/reaction})(1.602\times10^{-13} \text{ J/MeV})
\]  

(32.24)

\[
= 3.37\times10^{14} \text{ J.}
\]

**Solution for (b)**
Power is energy per unit time. One year has $3.16 \times 10^7$ s, so

$$P = \frac{E}{t} = \frac{3.37 \times 10^{14} \text{ J}}{3.16 \times 10^7 \text{ s}} = 1.07 \times 10^7 \text{ W} = 10.7 \text{ MW}.$$  

### Discussion

By now we expect nuclear processes to yield large amounts of energy, and we are not disappointed here. The energy output of $3.37 \times 10^{14} \text{ J}$ from fusing 1.00 kg of deuterium and tritium is equivalent to 2.6 million gallons of gasoline and about eight times the energy output of the bomb that destroyed Hiroshima. Yet the average backyard swimming pool has about 6 kg of deuterium in it, so that fuel is plentiful if it can be utilized in a controlled manner. The average power output over a year is more than 10 MW, impressive but a bit small for a commercial power plant. About 32 times this power output would allow generation of 100 MW of electricity, assuming an efficiency of one-third in converting the fusion energy to electrical energy.

### 32.6 Fission

**Nuclear fission** is a reaction in which a nucleus is split (or fissured). Controlled fission is a reality, whereas controlled fusion is a hope for the future. Hundreds of nuclear fission power plants around the world attest to the fact that controlled fission is practical and, at least in the short term, economical, as seen in Figure 32.24. Whereas nuclear power was of little interest for decades following TMI and Chernobyl (and now Fukushima Daiichi), growing concerns over global warming has brought nuclear power back on the table as a viable energy alternative. By the end of 2009, there were 442 reactors operating in 30 countries, providing 15% of the world’s electricity. France provides over 75% of its electricity with nuclear power, while the US has 104 operating reactors providing 20% of its electricity. Australia and New Zealand have none. China is building nuclear power plants at the rate of one start every month.

![Figure 32.24 The people living near this nuclear power plant have no measurable exposure to radiation that is traceable to the plant. About 16% of the world's electrical power is generated by controlled nuclear fission in such plants. The cooling towers are the most prominent features but are not unique to nuclear power. The reactor is in the small domed building to the left of the towers. (credit: Kalmthouts)](image)

Fission is the opposite of fusion and releases energy only when heavy nuclei are split. As noted in **Fusion**, energy is released if the products of a nuclear reaction have a greater binding energy per nucleon ($\frac{\text{BE}}{A}$) than the parent nuclei. Figure 32.25 shows that $\frac{\text{BE}}{A}$ is greater for medium-mass nuclei than heavy nuclei, implying that when a heavy nucleus is split, the products have less mass per nucleon, so that mass is destroyed and energy is released in the reaction. The amount of energy per fission reaction can be large, even by nuclear standards. The graph in Figure 32.25 shows $\frac{\text{BE}}{A}$ to be about 7.6 MeV/nucleon for the heaviest nuclei ($A$ about 240), while $\frac{\text{BE}}{A}$ is about 8.6 MeV/nucleon for nuclei having $A$ about 120. Thus, if a heavy nucleus splits in half, then about 1 MeV per nucleon, or approximately 240 MeV per fission, is released. This is about 10 times the energy per fusion reaction, and about 100 times the energy of the average $\alpha$, $\beta$, or $\gamma$ decay.

#### Example 32.3 Calculating Energy Released by Fission

Calculate the energy released in the following spontaneous fission reaction:

$$^{238}\text{U} \rightarrow ^{95}\text{Sr} + ^{140}\text{Xe} + 3n$$  

(32.26)

given the atomic masses to be $m(^{238}\text{U}) = 238.050784 \text{ u}$, $m(^{95}\text{Sr}) = 94.919388 \text{ u}$, $m(^{140}\text{Xe}) = 139.921610 \text{ u}$, and $m(n) = 1.008665 \text{ u}$.

**Strategy**

As always, the energy released is equal to the mass destroyed times $c^2$, so we must find the difference in mass between the parent $^{238}\text{U}$ and the fission products.

**Solution**

The products have a total mass of

$$m_{\text{total}} = m(^{95}\text{Sr}) + m(^{140}\text{Xe}) + 3m(n) = 139.921610 \text{ u} + 94.919388 \text{ u} + 3(1.008665 \text{ u}) = 229.859263 \text{ u}.$$  

The mass difference is

$$
\Delta m = m(^{238}\text{U}) - m_{\text{total}} = 238.050784 \text{ u} - 229.859263 \text{ u} = 8.191521 \text{ u}.
$$

The energy released is

$$E = \Delta m c^2 = 8.191521 \text{ u} \times (1.66054 \times 10^{-27} \text{ kg/u}) \times (3 \times 10^8 \text{ m/s})^2 = 3.62 \times 10^{14} \text{ J}.$$  

This content is available for free at http://cnx.org/content/col11406/1.7
\[ m_{\text{products}} = 94.919388 \text{ u} + 139.921610 \text{ u} + 3(1.008665 \text{ u}) \]
\[ = 237.866993 \text{ u}. \]  

The mass lost is the mass of \(^{238}\text{U}\) minus \(m_{\text{products}}\), or
\[ \Delta m = 238.050784 \text{ u} - 237.8669933 \text{ u} = 0.183791 \text{ u}, \]

so the energy released is
\[
E = (\Delta m)c^2 \\
= (0.183791 \text{ u}) \frac{931.5 \text{ MeV}}{\text{u}} c^2 = 171.2 \text{ MeV}. 
\]

**Discussion**

A number of important things arise in this example. The 171-MeV energy released is large, but a little less than the earlier estimated 240 MeV. This is because this fission reaction produces neutrons and does not split the nucleus into two equal parts. Fission of a given nuclide, such as \(^{238}\text{U}\), does not always produce the same products. Fission is a statistical process in which an entire range of products are produced with various probabilities. Most fission produces neutrons, although the number varies with each fission. This is an extremely important aspect of fission, because **neutrons can induce more fission**, enabling self-sustaining chain reactions.

Spontaneous fission can occur, but this is usually not the most common decay mode for a given nuclide. For example, \(^{238}\text{U}\) can spontaneously fission, but it decays mostly by \(\alpha\) emission. Neutron-induced fission is crucial as seen in [Figure 32.25](#). Being chargeless, even low-energy neutrons can strike a nucleus and be absorbed once they feel the attractive nuclear force. Large nuclei are described by a **liquid drop model** with surface tension and oscillation modes, because the large number of nucleons act like atoms in a drop. The neutron is attracted and thus, deposits energy, causing the nucleus to deform as a liquid drop. If stretched enough, the nucleus narrows in the middle. The number of nucleons in contact and the strength of the nuclear force binding the nucleus together are reduced. Coulomb repulsion between the two ends then succeeds in fissioning the nucleus, which pops like a water drop into two large pieces and a few neutrons. **Neutron-induced fission** can be written as
\[
n + ^{A}\text{X} \rightarrow FF_1 + FF_2 + xn, \quad (32.30)
\]

where \(FF_1\) and \(FF_2\) are the two daughter nuclei, called **fission fragments**, and \(x\) is the number of neutrons produced. Most often, the masses of the fission fragments are not the same. Most of the released energy goes into the kinetic energy of the fission fragments, with the remainder going into the neutrons and excited states of the fragments. Since neutrons can induce fission, a self-sustaining chain reaction is possible, provided more than one neutron is produced on average — that is, if \(x > 1\) in \(n + ^{A}\text{X} \rightarrow FF_1 + FF_2 + xn\). This can also be seen in [Figure 32.26](#).

An example of a typical neutron-induced fission reaction is
\[
n + ^{235}\text{U} \rightarrow ^{142}\text{Ba} + ^{91}\text{Kr} + 3n. \quad (32.31)
\]

Note that in this equation, the total charge remains the same (is conserved): \(92 + 0 = 56 + 36\). Also, as far as whole numbers are concerned, the mass is constant: \(1 + 235 = 142 + 91 + 3\). This is not true when we consider the masses out to 6 or 7 significant places, as in the previous example.
Neutron-induced fission is shown. First, energy is put into this large nucleus when it absorbs a neutron. Acting like a struck liquid drop, the nucleus deforms and begins to narrow in the middle. Since fewer nucleons are in contact, the repulsive Coulomb force is able to break the nucleus into two parts with some neutrons also flying away.

A chain reaction can produce self-sustained fission if each fission produces enough neutrons to induce at least one more fission. This depends on several factors, including how many neutrons are produced in an average fission and how easy it is to make a particular type of nuclide fission.

Not every neutron produced by fission induces fission. Some neutrons escape the fissionable material, while others interact with a nucleus without making it fission. We can enhance the number of fissions produced by neutrons by having a large amount of fissionable material. The minimum amount necessary for self-sustained fission of a given nuclide is called its **critical mass**. Some nuclides, such as $^{239\text{Pu}}$, produce more neutrons per fission than others, such as $^{235\text{U}}$. Additionally, some nuclides are easier to make fission than others. In particular, $^{235\text{U}}$ and $^{239\text{Pu}}$ are easier to fission than the much more abundant $^{238\text{U}}$. Both factors affect critical mass, which is smallest for $^{239\text{Pu}}$. 

---

**Figure 32.25** Neutron-induced fission is shown. First, energy is put into this large nucleus when it absorbs a neutron. Acting like a struck liquid drop, the nucleus deforms and begins to narrow in the middle. Since fewer nucleons are in contact, the repulsive Coulomb force is able to break the nucleus into two parts with some neutrons also flying away.

**Figure 32.26** A chain reaction can produce self-sustained fission if each fission produces enough neutrons to induce at least one more fission. This depends on several factors, including how many neutrons are produced in an average fission and how easy it is to make a particular type of nuclide fission.

Not every neutron produced by fission induces fission. Some neutrons escape the fissionable material, while others interact with a nucleus without making it fission. We can enhance the number of fissions produced by neutrons by having a large amount of fissionable material. The minimum amount necessary for self-sustained fission of a given nuclide is called its **critical mass**. Some nuclides, such as $^{239\text{Pu}}$, produce more neutrons per fission than others, such as $^{235\text{U}}$. Additionally, some nuclides are easier to make fission than others. In particular, $^{235\text{U}}$ and $^{239\text{Pu}}$ are easier to fission than the much more abundant $^{238\text{U}}$. Both factors affect critical mass, which is smallest for $^{239\text{Pu}}$. 

---

This content is available for free at http://cnx.org/content/col11406/1.7
The reason $^{235}\text{U}$ and $^{239}\text{Pu}$ are easier to fission than $^{238}\text{U}$ is that the nuclear force is more attractive for an even number of neutrons in a nucleus than for an odd number. Consider that $^{235}\text{U}_{143}$ has 143 neutrons, and $^{239}\text{Pu}_{145}$ has 145 neutrons, whereas $^{238}\text{U}_{146}$ has 146. When a neutron encounters a nucleus with an odd number of neutrons, the nuclear force is more attractive, because the additional neutron will make the number even. About 2-MeV more energy is deposited in the resulting nucleus than would be the case if the number of neutrons was already even. This extra energy produces greater deformation, making fission more likely. Thus, $^{235}\text{U}$ and $^{239}\text{Pu}$ are superior fission fuels. The isotope $^{235}\text{U}$ is only 0.72 % of natural uranium, while $^{238}\text{U}$ is 99.27%, and $^{239}\text{Pu}$ does not exist in nature. Australia has the largest deposits of uranium in the world, standing at 28% of the total. This is followed by Kazakhstan and Canada. The US has only 3% of global reserves.

Most fission reactors utilize $^{235}\text{U}$, which is separated from $^{238}\text{U}$ at some expense. This is called enrichment. The most common separation method is gaseous diffusion of uranium hexafluoride ($\text{UF}_6$) through membranes. Since $^{235}\text{U}$ has less mass than $^{238}\text{U}$, its $\text{UF}_6$ molecules have higher average velocity at the same temperature and diffuse faster. Another interesting characteristic of $^{235}\text{U}$ is that it preferentially absorbs very slow moving neutrons (with energies a fraction of an eV), whereas fission reactions produce fast neutrons with energies in the order of an MeV. To make a self-sustained fission reactor with $^{235}\text{U}$, it is thus necessary to slow down (“thermalize”) the neutrons. Water is very effective, since neutrons collide with protons in water molecules and lose energy.

Control rods containing nuclides that very strongly absorb neutrons are used to adjust neutron flux. To produce large power, reactors contain hundreds to thousands of critical masses, and the chain reaction easily becomes self-sustaining, a condition called criticality. Neutron flux should be carefully regulated to avoid an exponential increase in fissions, a condition called supercriticality. Control rods help prevent overheating, perhaps even a meltdown or explosive disassembly. The water that is used to thermalize neutrons, necessary to get them to induce fission in $^{235}\text{U}$, and achieve criticality, provides a negative feedback for temperature increases. In case the reactor overheats and boils the water away, the chain reaction terminates, because water is needed to thermalize the neutrons. This inherent safety feature can be overwhelmed in extreme circumstances.

**Example 32.4 Calculating Energy from a Kilogram of Fissionable Fuel**

Calculate the amount of energy produced by the fission of 1.00 kg of $^{235}\text{U}$, given the average fission reaction of $^{235}\text{U}$ produces 200 MeV.

**Strategy**

The total energy produced is the number of $^{235}\text{U}$ atoms times the given energy per $^{235}\text{U}$ fission. We should therefore find the number of $^{235}\text{U}$ atoms in 1.00 kg.

**Solution**

Figure 32.27 A pressurized water reactor is cleverly designed to control the fission of large amounts of $^{235}\text{U}$, while using the heat produced in the fission reaction to create steam for generating electrical energy. Control rods adjust neutron flux so that criticality is obtained, but not exceeded. In case the reactor overheats and boils the water away, the chain reaction terminates, because water is needed to thermalize the neutrons. This inherent safety feature can be overwhelmed in extreme circumstances.
The number of $^{235}$U atoms in 1.00 kg is Avogadro's number times the number of moles. One mole of $^{235}$U has a mass of 235.04 g; thus, there are $(1000 \text{ g}) / (235.04 \text{ g/mol}) = 4.25 \text{ mol}$. The number of $^{235}$U atoms is therefore,

$$(4.25 \text{ mol})(6.02 \times 10^{23} \text{ }^{235}\text{U/mol}) = 2.56 \times 10^{24} \text{ }^{235}\text{U}.$$  

So the total energy released is

$$E = (2.56 \times 10^{24} \text{ }^{235}\text{U}) \left( \frac{200 \text{ MeV}}{235 \text{ U}} \right) \left( \frac{1.60 \times 10^{-13} \text{ J}}{\text{MeV}} \right)$$

$$= 8.21 \times 10^{13} \text{ J}.$$  

**Discussion**

This is another impressively large amount of energy, equivalent to about 14,000 barrels of crude oil or 600,000 gallons of gasoline. But, it is only one-fourth the energy produced by the fusion of a kilogram mixture of deuterium and tritium as seen in Example 32.2. Even though each fission reaction yields about ten times the energy of a fusion reaction, the energy per kilogram of fission fuel is less, because there are far fewer moles per kilogram of the heavy nuclides. Fission fuel is also much more scarce than fusion fuel, and less than 1% of uranium (the $^{235}$U) is readily usable.

One nuclide already mentioned is $^{239}$Pu, which has a 24,120-y half-life and does not exist in nature. Plutonium-239 is manufactured from $^{238}$U in reactors, and it provides an opportunity to utilize the other 99% of natural uranium as an energy source. The following reaction sequence, called **breeding**, produces $^{239}$Pu. Breeding begins with neutron capture by $^{238}$U:

$$^{238}\text{U} + n \rightarrow ^{239}\text{U} + \gamma.$$  

Uranium-239 then $\beta^-$ decays:

$$^{239}\text{U} \rightarrow ^{239}\text{Np} + \beta^- + v_e (t_{1/2} = 23 \text{ min}).$$  

Neptunium-239 also $\beta^-$ decays:

$$^{239}\text{Np} \rightarrow ^{239}\text{Pu} + \beta^- + v_e (t_{1/2} = 2.4 \text{ d}).$$

Plutonium-239 builds up in reactor fuel at a rate that depends on the probability of neutron capture by $^{238}$U (all reactor fuel contains more $^{238}$U than $^{235}$U). Reactors designed specifically to make plutonium are called **breeder reactors**. They seem to be inherently more hazardous than conventional reactors, but it remains unknown whether their hazards can be made economically acceptable. The four reactors at Chernobyl, including the one that was destroyed, were built to breed plutonium and produce electricity. These reactors had a design that was significantly different from the pressurized water reactor illustrated above.

Plutonium-239 has advantages over $^{235}$U as a reactor fuel — it produces more neutrons per fission on average, and it is easier for a thermal neutron to cause it to fission. It is also chemically different from uranium, so it is inherently easier to separate from uranium ore. This means $^{239}$Pu has a particularly small critical mass, an advantage for nuclear weapons.

**PhET Explorations: Nuclear Fission**

Start a chain reaction, or introduce non-radioactive isotopes to prevent one. Control energy production in a nuclear reactor!

**PhET Interactive Simulation**

Figure 32.28 Nuclear Fission (http://cnx.org/content/m42662/1.7/nuclear-fission_en.jar)

### 32.7 Nuclear Weapons

The world was in turmoil when fission was discovered in 1938. The discovery of fission, made by two German physicists, Otto Hahn and Fritz Strassman, was quickly verified by two Jewish refugees from Nazi Germany, Lise Meitner and her nephew Otto Frisch. Fermi, among others, soon found that not only did neutrons induce fission; more neutrons were produced during fission. The possibility of a self-sustained chain reaction was immediately recognized by leading scientists the world over. The enormous energy known to be in nuclei, but considered inaccessible, now seemed to be available on a large scale.
Within months after the announcement of the discovery of fission, Adolf Hitler banned the export of uranium from newly occupied Czechoslovakia. It seemed that the military value of uranium had been recognized in Nazi Germany, and that a serious effort to build a nuclear bomb had begun.

Alarmed scientists, many of them who fled Nazi Germany, decided to take action. None was more famous or revered than Einstein. It was felt that his help was needed to get the American government to make a serious effort at nuclear weapons as a matter of survival. Leo Szilard, an escaped Hungarian physicist, took a draft of a letter to Einstein, who, although pacifistic, signed the final version. The letter was for President Franklin Roosevelt, warning of the German potential to build extremely powerful bombs of a new type. It was sent in August of 1939, just before the German invasion of Poland that marked the start of World War II.

It was not until December 6, 1941, the day before the Japanese attack on Pearl Harbor, that the United States made a massive commitment to building a nuclear bomb. The top secret Manhattan Project was a crash program aimed at beating the Germans. It was carried out in remote locations, such as Los Alamos, New Mexico, whenever possible, and eventually came to cost billions of dollars and employ the efforts of more than 100,000 people. J. Robert Oppenheimer (1904–1967), whose talent and ambitions made him ideal, was chosen to head the project. The first major step was made by Enrico Fermi and his group in December 1942, when they achieved the first self-sustained nuclear reactor. This first “atomic pile”, built in a squash court at the University of Chicago, used carbon blocks to thermalize neutrons. It not only proved that the chain reaction was possible, it began the era of nuclear reactors. Glenn Seaborg, an American chemist and physicist, received the Nobel Prize in physics in 1951 for discovery of several transuranic elements, including plutonium. Carbon-moderated reactors are relatively inexpensive and simple in design and are still used for breeding plutonium, such as at Chernobyl, where two such reactors remain in operation.

Plutonium was recognized as easier to fission with neutrons and, hence, a superior fission material very early in the Manhattan Project. Plutonium availability was uncertain, and so a uranium bomb was developed simultaneously. Figure 32.29 shows a gun-type bomb, which takes two subcritical uranium masses and blows them together. To get an appreciable yield, the critical mass must be held together by the explosive charges inside the cannon barrel for a few microseconds. Since the buildup of the uranium chain reaction is relatively slow, the device to hold the critical mass together can be relatively simple. Owing to the fact that the rate of spontaneous fission is low, a neutron source is triggered at the same time the critical mass is assembled.

![Figure 32.29](image)

A gun-type fission bomb for $^{235}\text{U}$ utilizes two subcritical masses forced together by explosive charges inside a cannon barrel. The energy yield depends on the amount of uranium and the time it can be held together before it disassembles itself.

Plutonium's special properties necessitated a more sophisticated critical mass assembly, shown schematically in Figure 32.30. A spherical mass of plutonium is surrounded by shape charges (high explosives that release most of their blast in one direction) that implode the plutonium, crushing it into a smaller volume to form a critical mass. The implosion technique is faster and more effective, because it compresses three-dimensionally rather than one-dimensionally as in the gun-type bomb. Again, a neutron source must be triggered at just the correct time to initiate the chain reaction.
An implosion created by high explosives compresses a sphere of $^{239}\text{Pu}$ into a critical mass. The superior fissionability of plutonium has made it the universal bomb material.

Owing to its complexity, the plutonium bomb needed to be tested before there could be any attempt to use it. On July 16, 1945, the test named Trinity was conducted in the isolated Alamogordo Desert about 200 miles south of Los Alamos (see Figure 32.31). A new age had begun. The yield of this device was about 10 kilotons (kT), the equivalent of 5000 of the largest conventional bombs.

Figure 32.31  Trinity test (1945), the first nuclear bomb (credit: United States Department of Energy)

Although Germany surrendered on May 7, 1945, Japan had been steadfastly refusing to surrender for many months, forcing large casualties. Invasion plans by the Allies estimated a million casualties of their own and untold losses of Japanese lives. The bomb was viewed as a way to end the war. The first was a uranium bomb dropped on Hiroshima on August 6. Its yield of about 15 kT destroyed the city and killed an estimated 80,000 people, with 100,000 more being seriously injured (see Figure 32.32). The second was a plutonium bomb dropped on Nagasaki only three days later, on August 9. Its 20 kT yield killed at least 50,000 people, something less than Hiroshima because of the hilly terrain and the fact that it was a few kilometers off target. The Japanese were told that one bomb a week would be dropped until they surrendered unconditionally, which they did on August 14. In actuality, the United States had only enough plutonium for one more and as yet unassembled bomb.

Knowing that fusion produces several times more energy per kilogram of fuel than fission, some scientists pushed the idea of a fusion bomb starting very early on. Calling this bomb the Super, they realized that it could have another advantage over fission—high-energy neutrons would aid fusion, while they are ineffective in $^{239}\text{Pu}$ fission. Thus the fusion bomb could be virtually unlimited in energy release. The first such bomb was detonated...
by the United States on October 31, 1952, at Eniwetok Atoll with a yield of 10 megatons (MT), about 670 times that of the fission bomb that destroyed Hiroshima. The Soviets followed with a fusion device of their own in August 1953, and a weapons race, beyond the aim of this text to discuss, continued until the end of the Cold War.

**Figure 32.33** shows a simple diagram of how a thermonuclear bomb is constructed. A fission bomb is exploded next to fusion fuel in the solid form of lithium deuteride. Before the shock wave blows it apart, $\gamma$ rays heat and compress the fuel, and neutrons create tritium through the reaction $n + ^{6}\text{Li} \rightarrow ^{3}\text{H} + ^{4}\text{He}$. Additional fusion and fission fuels are enclosed in a dense shell of $^{238}\text{U}$. The shell reflects some of the neutrons back into the fuel to enhance its fusion, but at high internal temperatures fast neutrons are created that also cause the plentiful and inexpensive $^{238}\text{U}$ to fission, part of what allows thermonuclear bombs to be so large.

---

**Figure 32.33** This schematic of a fusion bomb (H-bomb) gives some idea of how the $^{239}\text{Pu}$ fission trigger is used to ignite fusion fuel. Neutrons and $\gamma$ rays transmit energy to the fusion fuel, create tritium from deuterium, and heat and compress the fusion fuel. The outer shell of $^{238}\text{U}$ serves to reflect some neutrons back into the fuel, causing more fusion, and it boosts the energy output by fissioning itself when neutron energies become high enough.

The energy yield and the types of energy produced by nuclear bombs can be varied. Energy yields in current arsenals range from about 0.1 kT to 20 MT, although the Soviets once detonated a 67 MT device. Nuclear bombs differ from conventional explosives in more than size. **Figure 32.34** shows the approximate fraction of energy output in various forms for conventional explosives and for two types of nuclear bombs. Nuclear bombs put a much larger fraction of their output into thermal energy than do conventional bombs, which tend to concentrate the energy in blast. Another difference is the immediate and residual radiation energy from nuclear weapons. This can be adjusted to put more energy into radiation (the so-called neutron bomb) so that the bomb can be used to irradiate advancing troops without killing friendly troops with blast and heat.
At its peak in 1986, the combined arsenals of the United States and the Soviet Union totaled about 60,000 nuclear warheads. In addition, the British, French, and Chinese each have several hundred bombs of various sizes, and a few other countries have a small number. Nuclear weapons are generally divided into two categories. Strategic nuclear weapons are those intended for military targets, such as bases and missile complexes, and moderate to large cities. There were about 20,000 strategic weapons in 1988. Tactical weapons are intended for use in smaller battles. Since the collapse of the Soviet Union and the end of the Cold War in 1989, most of the 32,000 tactical weapons (including Cruise missiles, artillery shells, land mines, torpedoes, depth charges, and backpacks) have been demobilized, and parts of the strategic weapon systems are being dismantled with warheads and missiles being disassembled. According to the Treaty of Moscow of 2002, Russia and the United States have been required to reduce their strategic nuclear arsenal down to about 2000 warheads each.

A few small countries have built or are capable of building nuclear bombs, as are some terrorist groups. Two things are needed—a minimum level of technical expertise and sufficient fissionable material. The first is easy. Fissionable material is controlled but is also available. There are international agreements and organizations that attempt to control nuclear proliferation, but it is increasingly difficult given the availability of fissionable material and the small amount needed for a crude bomb. The production of fissionable fuel itself is technologically difficult. However, the presence of large amounts of such material worldwide, though in the hands of a few, makes control and accountability crucial.

**Glossary**

**Anger camera**: a common medical imaging device that uses a scintillator connected to a series of photomultipliers

**break-even**: when fusion power produced equals the heating power input

**breeder reactors**: reactors that are designed specifically to make plutonium

**breeding**: reaction process that produces $^{239}$Pu

**critical mass**: minimum amount necessary for self-sustained fission of a given nuclide

**criticality**: condition in which a chain reaction easily becomes self-sustaining
fission fragments: a daughter nuclei

food irradiation: treatment of food with ionizing radiation

free radicals: ions with unstable oxygen- or hydrogen-containing molecules

gamma camera: another name for an Anger camera

gray (Gy): the SI unit for radiation dose which is defined to be 1 Gy = 1 J/kg = 100 rad

high dose: a dose greater than 1 Sv (100 rem)

hormesis: a term used to describe generally favorable biological responses to low exposures of toxins or radiation

ignition: when a fusion reaction produces enough energy to be self-sustaining after external energy input is cut off

inertial confinement: a technique that aims multiple lasers at tiny fuel pellets evaporating and crushing them to high density

linear hypothesis: assumption that risk is directly proportional to risk from high doses

liquid drop model: a model of nucleus (only to understand some of its features) in which nucleons in a nucleus act like atoms in a drop

low dose: a dose less than 100 mSv (10 rem)

magnetic confinement: a technique in which charged particles are trapped in a small region because of difficulty in crossing magnetic field lines

moderate dose: a dose from 0.1 Sv to 1 Sv (10 to 100 rem)

neutron-induced fission: fission that is initiated after the absorption of neutron

nuclear fission: reaction in which a nucleus splits

nuclear fusion: a reaction in which two nuclei are combined, or fused, to form a larger nucleus

positron emission tomography (PET): tomography technique that uses $\beta^+$ emitters and detects the two annihilation $\gamma$ rays, aiding in source localization

proton-proton cycle: the combined reactions $^1\text{H}+^1\text{H} \rightarrow ^2\text{He}+e^++\nu_e$, $^1\text{H}+^2\text{H} \rightarrow ^3\text{He}+\gamma$, and $^3\text{He}+^3\text{He} \rightarrow ^4\text{He}+^4\text{He}+^1\text{H}$

quality factor: same as relative biological effectiveness

radiolytic products: compounds produced due to chemical reactions of free radicals

radiopharmaceutical: compound used for medical imaging

radiotherapy: the use of ionizing radiation to treat ailments

rad: the ionizing energy deposited per kilogram of tissue

relative biological effectiveness (RBE): a number that expresses the relative amount of damage that a fixed amount of ionizing radiation of a given type can inflict on biological tissues

roentgen equivalent man (rem): a dose unit more closely related to effects in biological tissue

shielding: a technique to limit radiation exposure

sievert: the SI equivalent of the rem

single-photon-emission computed tomography (SPECT): tomography performed with $\gamma$-emitting radiopharmaceuticals

supercriticality: an exponential increase in fissions

tagged: process of attaching a radioactive substance to a chemical compound

therapeutic ratio: the ratio of abnormal cells killed to normal cells killed

Section Summary

32.1 Medical Imaging and Diagnostics
- Radiopharmaceuticals are compounds that are used for medical imaging and therapeutics.
- The process of attaching a radioactive substance is called tagging.
- Table 32.1 lists certain diagnostic uses of radiopharmaceuticals including the isotope and activity typically used in diagnostics.
- One common imaging device is the Anger camera, which consists of a lead collimator, radiation detectors, and an analysis computer.
- Tomography performed with $\gamma$-emitting radiopharmaceuticals is called SPECT and has the advantages of x-ray CT scans coupled with organ- and function-specific drugs.
32.2 Biological Effects of Ionizing Radiation
- The biological effects of ionizing radiation are due to two effects it has on cells: interference with cell reproduction, and destruction of cell function.
- A radiation dose unit called the rad is defined in terms of the ionizing energy deposited per kilogram of tissue:
  \[ 1 \text{ rad} = 0.01 \text{ J/kg}. \]
- The SI unit for radiation dose is the gray (Gy), which is defined to be \( 1 \text{ Gy} = 1 \text{ J/kg} = 100 \text{ rad} \).
- To account for the effect of the type of particle creating the ionization, we use the relative biological effectiveness (RBE) or quality factor (QF) given in Table 32.2 and define a unit called the roentgen equivalent man (rem) as
  \[ \text{rem} = \text{rad} \times \text{RBE}. \]
- Particles that have short ranges or create large ionization densities have RBEs greater than unity. The SI equivalent of the rem is the sievert (Sv), defined to be
  \[ \text{Sv} = \text{Gy} \times \text{RBE} \text{ and } 1 \text{ Sv} = 100 \text{ rem}. \]
- Whole-body, single-exposure doses of 0.1 Sv or less are low doses while those of 0.1 to 1 Sv are moderate, and those over 1 Sv are high doses. Some immediate radiation effects are given in Table 32.4. Effects due to low doses are not observed, but their risk is assumed to be directly proportional to those of high doses, an assumption known as the linear hypothesis. Long-term effects are cancer deaths at the rate of \( 10 / 10^6 \) rem·y and genetic defects at roughly one-third this rate. Background radiation doses and sources are given in Table 32.5. Worldwide average radiation exposure from natural sources, including radon, is about 3 mSv, or 300 mrem. Radiation protection utilizes shielding, distance, and time to limit exposure.

32.3 Therapeutic Uses of Ionizing Radiation
- Radiotherapy is the use of ionizing radiation to treat ailments, now limited to cancer therapy.
- The sensitivity of cancer cells to radiation enhances the ratio of cancer cells killed to normal cells killed, which is called the therapeutic ratio.
- Doses for various organs are limited by the tolerance of normal tissue for radiation. Treatment is localized in one region of the body and spread out in time.

32.5 Fusion
- Nuclear fusion is a reaction in which two nuclei are combined to form a larger nucleus. It releases energy when light nuclei are fused to form medium-mass nuclei.
- Fusion is the source of energy in stars, with the proton-proton cycle,
  \[ \begin{align*} \text{H} + \text{H} & \rightarrow \text{H} + \text{He} + e^+ + \nu_e \quad (0.42 \text{ MeV}) \\ \text{H} + \text{H} & \rightarrow \text{He} + \gamma \quad (5.49 \text{ MeV}) \end{align*} \]
- being the principal sequence of energy-producing reactions in our Sun.
- The overall effect of the proton-proton cycle is
  \[ 2e^- + \text{H} \rightarrow \text{He} + 2\nu_e + 6\gamma \quad (26.7 \text{ MeV}), \]
  where the 26.7 MeV includes the energy of the positrons emitted and annihilated.
- Attempts to utilize controlled fusion as an energy source on Earth are related to deuterium and tritium, and the reactions play important roles.
- Ignition is the condition under which controlled fusion is self-sustaining; it has not yet been achieved. Break-even, in which the fusion energy output is as great as the external energy input, has nearly been achieved.
- Magnetic confinement and inertial confinement are the two methods being developed for heating fuel to sufficiently high temperatures, at sufficient density, and for sufficiently long times to achieve ignition. The first method uses magnetic fields and the second method uses the momentum of impinging laser beams for confinement.

32.6 Fission
- Nuclear fission is a reaction in which a nucleus is split.
- Fission releases energy when heavy nuclei are split into medium-mass nuclei.
- Self-sustained fission is possible, because neutron-induced fission also produces neutrons that can induce other fissions,
  \[ n + X \rightarrow FF_1 + FF_2 + x\gamma, \text{ where } FF_1 \text{ and } FF_2 \text{ are the two daughter nuclei, or fission fragments, and } x \text{ is the number of neutrons produced}. \]
- A minimum mass, called the critical mass, should be present to achieve criticality.
- More than a critical mass can produce supercriticality.
- The production of new or different isotopes (especially \( ^{239}\text{Pu} \)) by nuclear transformation is called breeding, and reactors designed for this purpose are called breeder reactors.

32.7 Nuclear Weapons
- There are two types of nuclear weapons—fission bombs use fission alone, whereas thermonuclear bombs use fission to ignite fusion.
- Both types of weapons produce huge numbers of nuclear reactions in a very short time.
- Energy yields are measured in kilotons or megatons of equivalent conventional explosives and range from 0.1 kT to more than 20 MT.
- Nuclear bombs are characterized by far more thermal output and nuclear radiation output than conventional explosives.
Conceptual Questions

32.1 Medical Imaging and Diagnostics

1. In terms of radiation dose, what is the major difference between medical diagnostic uses of radiation and medical therapeutic uses?

2. One of the methods used to limit radiation dose to the patient in medical imaging is to employ isotopes with short half-lives. How would this limit the dose?

32.2 Biological Effects of Ionizing Radiation

3. Isotopes that emit $\alpha$ radiation are relatively safe outside the body and exceptionally hazardous inside. Yet those that emit $\gamma$ radiation are hazardous outside and inside. Explain why.

4. Why is radon more closely associated with inducing lung cancer than other types of cancer?

5. The RBE for low-energy $\beta$ s is 1.7, whereas that for higher-energy $\beta$ s is only 1. Explain why, considering how the range of radiation depends on its energy.

6. Which methods of radiation protection were used in the device shown in the first photo in Figure 32.35? Which were used in the situation shown in the second photo?

(a)

Figure 32.35 (a) This x-ray fluorescence machine is one of the thousands used in shoe stores to produce images of feet as a check on the fit of shoes. They are unshielded and remain on as long as the feet are in them, producing doses much greater than medical images. Children were fascinated with them. These machines were used in shoe stores until laws preventing such unwarranted radiation exposure were enacted in the 1950s. (credit: Andrew Kuchling ) (b) Now that we know the effects of exposure to radioactive material, safety is a priority. (credit: U.S. Navy)

7. What radioisotope could be a problem in homes built of cinder blocks made from uranium mine tailings? (This is true of homes and schools in certain regions near uranium mines.)

8. Are some types of cancer more sensitive to radiation than others? If so, what makes them more sensitive?

9. Suppose a person swallows some radioactive material by accident. What information is needed to be able to assess possible damage?

32.3 Therapeutic Uses of Ionizing Radiation

10. Radiotherapy is more likely to be used to treat cancer in elderly patients than in young ones. Explain why. Why is radiotherapy used to treat young people at all?

32.4 Food Irradiation

11. Does food irradiation leave the food radioactive? To what extent is the food altered chemically for low and high doses in food irradiation?

12. Compare a low dose of radiation to a human with a low dose of radiation used in food treatment.

13. Suppose one food irradiation plant uses a $^{137}\text{Cs}$ source while another uses an equal activity of $^{60}\text{Co}$. Assuming equal fractions of the $\gamma$ rays from the sources are absorbed, why is more time needed to get the same dose using the $^{137}\text{Cs}$ source?

32.5 Fusion
14. Why does the fusion of light nuclei into heavier nuclei release energy?

15. Energy input is required to fuse medium-mass nuclei, such as iron or cobalt, into more massive nuclei. Explain why.

16. In considering potential fusion reactions, what is the advantage of the reaction \( ^2\text{H} + ^3\text{H} \rightarrow ^4\text{He} + n \) over the reaction \( ^2\text{H} + ^2\text{H} \rightarrow ^3\text{He} + n \)?

17. Give reasons justifying the contention made in the text that energy from the fusion reaction \( ^2\text{H} + ^2\text{H} \rightarrow ^4\text{He} + \gamma \) is relatively difficult to capture and utilize.

### 32.6 Fission

18. Explain why the fission of heavy nuclei releases energy. Similarly, why is it that energy input is required to fission light nuclei?

19. Explain, in terms of conservation of momentum and energy, why collisions of neutrons with protons will thermalize neutrons better than collisions with oxygen.

20. The ruins of the Chernobyl reactor are enclosed in a huge concrete structure built around it after the accident. Some rain penetrates the building in winter, and radioactivity from the building increases. What does this imply is happening inside?

21. Since the uranium or plutonium nucleus fissions into several fission fragments whose mass distribution covers a wide range of pieces, would you expect more residual radioactivity from fission than fusion? Explain.

22. The core of a nuclear reactor generates a large amount of thermal energy from the decay of fission products, even when the power-producing fission chain reaction is turned off. Would this residual heat be greatest after the reactor has run for a long time or short time? What if the reactor has been shut down for months?

23. How can a nuclear reactor contain many critical masses and not go supercritical? What methods are used to control the fission in the reactor?

24. Why can heavy nuclei with odd numbers of neutrons be induced to fission with thermal neutrons, whereas those with even numbers of neutrons require more energy input to induce fission?

25. Why is a conventional fission nuclear reactor not able to explode as a bomb?

### 32.7 Nuclear Weapons

26. What are some of the reasons that plutonium rather than uranium is used in all fission bombs and as the trigger in all fusion bombs?

27. Use the laws of conservation of momentum and energy to explain how a shape charge can direct most of the energy released in an explosion in a specific direction. (Note that this is similar to the situation in guns and cannons—most of the energy goes into the bullet.)

28. How does the lithium deuteride in the thermonuclear bomb shown in Figure 32.33 supply tritium \(^3\text{H}\) as well as deuterium \(^2\text{H}\)?

29. Fallout from nuclear weapons tests in the atmosphere is mainly \(^{90}\text{Sr}\) and \(^{137}\text{Cs}\), which have 28.6- and 32.2-y half-lives, respectively. Atmospheric tests were terminated in most countries in 1963, although China only did so in 1980. It has been found that environmental activities of these two isotopes are decreasing faster than their half-lives. Why might this be?
32.1 Medical Imaging and Diagnostics

1. A neutron generator uses an $\alpha$ source, such as radium, to bombard beryllium, inducing the reaction $^4\text{He} + ^9\text{Be} \rightarrow ^{12}\text{C} + n$. Such neutron sources are called RaBe sources, or PuBe sources if they use plutonium to get the $\alpha$ s. Calculate the energy output of the reaction in MeV.

2. Neutrons from a source (perhaps the one discussed in the preceding problem) bombard natural molybdenum, which is 24 percent $^{98}\text{Mo}$. What is the energy output of the reaction $^{98}\text{Mo} + n \rightarrow ^{99}\text{Mo} + \gamma$? The mass of $^{98}\text{Mo}$ is given in Appendix A: Atomic Masses, and that of $^{99}\text{Mo}$ is 98.907711 u.

3. The purpose of producing $^{99}\text{Mo}$ (usually by neutron activation of natural molybdenum, as in the preceding problem) is to produce $^{99m}\text{Tc}$. Using the rules, verify that the $\beta^-$ decay of $^{99}\text{Mo}$ produces $^{99m}\text{Tc}$. (Most $^{99m}\text{Tc}$ nuclei produced in this decay are left in a metastable excited state denoted $^{99m}\text{Tc^*}$).

4. (a) Two annihilation $\gamma$ rays in a PET scan originate at the same point and travel to detectors on either side of the patient. If the point of origin is 9.00 cm closer to one of the detectors, what is the difference in arrival times of the photons? (This could be used to give position information, but the time difference is small enough to make it difficult.)

(b) How accurately would you need to be able to measure arrival time differences to get a position resolution of 1.00 mm?

5. Table 32.1 indicates that 7.50 mCi of $^{99m}\text{Tc}$ is used in a brain scan. What is the mass of technetium?

6. The activities of $^{131}\text{I}$ and $^{123}\text{I}$ used in thyroid scans are given in Table 32.1 to be 50 and $70$ µCi, respectively. Find and compare the masses of $^{131}\text{I}$ and $^{123}\text{I}$ in such scans, given their respective half-lives are 8.04 d and 13.2 h. The masses are so small that the radiiodine is usually mixed with stable iodine as a carrier to ensure normal chemistry and distribution in the body.

7. (a) Neutron activation of sodium, which is 100% $^{23}\text{Na}$, produces $^{24}\text{Na}$, which is used in some heart scans, as seen in Table 32.1. The equation for the reaction is $^{23}\text{Na} + n \rightarrow ^{24}\text{Na} + \gamma$. Find its energy output, given the mass of $^{24}\text{Na}$ is 23.990962 u.

(b) What mass of $^{24}\text{Na}$ produces the needed 5.0-mCi activity, given its half-life is 15.0 h?

32.2 Biological Effects of Ionizing Radiation

8. What is the dose in mSv for: (a) a 0.1 Gy x-ray? (b) 2.5 mGy of neutron exposure to the eye? (c) 1.5 mGy of $\alpha$ exposure?

9. Find the radiation dose in Gy for: (a) A 10-mSv fluoroscopic x-ray series. (b) 50 mSv of skin exposure by an $\alpha$ emitter. (c) 160 mSv of $\beta^-$ and $\gamma$ rays from the $^{40}\text{K}$ in your body.

10. How many Gy of exposure is needed to give a cancerous tumor a dose of 40 Sv if it is exposed to $\alpha$ activity?

11. What is the dose in Sv in a cancer treatment that exposes the patient to 200 Gy of $\gamma$ rays?

12. One half the $\gamma$ rays from $^{99m}\text{Tc}$ are absorbed by a 0.170-mm-thick lead shielding. Half of the $\gamma$ rays that pass through the first layer of lead are absorbed in a second layer of equal thickness. What thickness of lead will absorb all but one in 1000 of these $\gamma$ rays?

13. A plumber at a nuclear power plant receives a whole-body dose of 30 mSv in 15 minutes while repairing a crucial valve. Find the radiation-induced yearly risk of death from cancer and the chance of genetic defect from this maximum allowable exposure.

14. In the 1980s, the term picowave was used to describe food irradiation in order to overcome public resistance by playing on the well-known safety of microwave radiation. Find the energy in MeV of a photon having a wavelength of a picometer.

15. Find the mass of $^{239}\text{Pu}$ that has an activity of 1.00 µCi.

32.3 Therapeutic Uses of Ionizing Radiation

16. A beam of 168-MeV nitrogen nuclei is used for cancer therapy. If this beam is directed onto a 0.200-kg tumor and gives it a 2.00-Sv dose, how many nitrogen nuclei were stopped? (Use an RBE of 20 for heavy ions.)

17. (a) If the average molecular mass of compounds in food is 50.0 g, how many molecules are there in 1.00 kg of food? (b) How many ion pairs are created in 1.00 kg of food, if it is exposed to 1000 Sv and it takes 32.0 eV to create an ion pair? (c) Find the ratio of ion pairs to molecules. (d) If these ion pairs recombine into a distribution of 2000 new compounds, how many parts per billion is each?

18. Calculate the dose in Sv to the chest of a patient given an x-ray under the following conditions. The x-ray beam intensity is 1.50 W/m², the area of the chest exposed is 0.0750 m², 35.0% of the x-rays are absorbed in 20.0 kg of tissue, and the exposure time is 0.250 s.

19. (a) A cancer patient is exposed to $\gamma$ rays from a 5000-Ci $^{60}\text{Co}$ transillumination unit for 32.0 s. The $\gamma$ rays are collimated in such a manner that only 1.00% of them strike the patient. Of those, 20.0% are absorbed in a tumor having a mass of 1.50 kg. What is the dose in rem to the tumor, if the average $\gamma$ energy per decay is 1.25 MeV? None of the $\beta$ s from the decay reach the patient. (b) Is the dose consistent with stated therapeutic doses?

20. What is the mass of $^{60}\text{Co}$ in a cancer therapy transillumination unit containing 5.00 kCi of $^{60}\text{Co}$?

21. Large amounts of $^{65}\text{Zn}$ are produced in copper exposed to accelerator beams. While machining contaminated copper, a physicist ingests 50.0 µCi of $^{65}\text{Zn}$. Each $^{65}\text{Zn}$ decay emits an average $\gamma$-ray energy of 0.550 MeV, 40.0% of which is absorbed in the scientist’s 75.0-kg body. What dose in mSv is caused by this in one day?

22. Naturally occurring $^{40}\text{K}$ is listed as responsible for 16 mrem/y of background radiation. Calculate the mass of $^{40}\text{K}$ that must be inside the 55-kg body of a woman to produce this dose. Each $^{40}\text{K}$ decay emits a 1.32-MeV $\beta^-$, and 50% of the energy is absorbed inside the body.

23. (a) Background radiation due to $^{226}\text{Ra}$ averages only 0.01 mSv/y, but it can range upward depending on where a person lives. Find the mass of $^{226}\text{Ra}$ in the 80.0-kg body of a man who receives a dose of 2.50-mSv/y from it, noting that each $^{226}\text{Ra}$ decay emits a 4.80-MeV $\alpha$ particle. You may neglect dose due to daughters and assume a constant amount, evenly distributed due to balanced ingestion and bodily
elimination. (b) Is it surprising that such a small mass could cause a measurable radiation dose? Explain.

24. The annual radiation dose from $^{14}$C in our bodies is 0.01 mSv/y. Each $^{14}$C decay emits a $\beta^-$ averaging 0.0750 MeV. Taking the fraction of $^{14}$C to be $1.3 \times 10^{-12}$ N of normal $^{12}$C, and assuming the body is 13% carbon, estimate the fraction of the decay energy absorbed. (The rest escapes, exposing those close to you.) Assume that there are $200 \times 10^{-4}$ deaths per Sv of radiation per year. What percent of the actual number of cancer deaths recorded is this?

32.5 Fusion

26. Verify that the total number of nucleons, total charge, and electron family number are conserved for each of the fusion reactions in the proton-proton cycle in

$$^{1}\text{H} + ^{1}\text{H} \rightarrow ^{2}\text{H} + e^+ + v_e, \quad ^{1}\text{H} + ^{2}\text{H} \rightarrow ^{3}\text{He} + \gamma,$$

and

$$^{3}\text{He} + ^{3}\text{He} \rightarrow ^{4}\text{He} + ^{1}\text{H} + ^{1}\text{H}.$$ 

(List the value of each of the conserved quantities before and after each of the reactions.)

27. Calculate the energy output in each of the fusion reactions in the proton-proton cycle, and verify the values given in the above summary.

28. Show that the total energy released in the proton-proton cycle is 26.7 MeV, considering the overall effect in $^{1}\text{H} + ^{1}\text{H} \rightarrow ^{2}\text{H} + e^+ + v_e.$

29. Verify by listing the number of nucleons, total charge, and electron family number before and after the cycle that these quantities are conserved in the overall proton-proton cycle in $2e^- + ^{4}\text{He} \rightarrow ^{4}\text{He} + 2v_e + 6\gamma.$

30. The energy produced by the fusion of a 1.00-kg mixture of deuterium and tritium was found in Example Calculating Energy and Power from Fusion. Approximately how many kilograms would be required to supply the annual energy use in the United States?

31. Tritium is naturally rare, but can be produced by the reaction $n + ^{2}\text{H} \rightarrow ^{3}\text{H} + \gamma.$ How much energy in MeV is released in this neutron capture?

32. Two fusion reactions mentioned in the text are

$$n + ^{3}\text{He} \rightarrow ^{4}\text{He} + \gamma$$

and

$$n + ^{1}\text{H} \rightarrow ^{2}\text{H} + \gamma.$$ 

Both reactions release energy, but the second also creates more fuel. Confirm that the energies produced in the reactions are 20.58 and 2.22 MeV, respectively. Comment on which product nuclide is most tightly bound, $^{4}\text{He}$ or $^{2}\text{H}.$

33. (a) Calculate the number of grams of deuterium in an 80,000-L swimming pool, given deuterium is 0.0150% of natural hydrogen.

(b) Find the energy released in joules if this deuterium is fused via the reaction $^{2}\text{H} + ^{2}\text{H} \rightarrow ^{3}\text{He} + n.$

(c) Could the neutrons be used to create more energy?

(d) Discuss the amount of this type of energy in a swimming pool as compared to that in, say, a gallon of gasoline, also taking into consideration that water is far more abundant.

34. How many kilograms of water are needed to obtain the 198.8 mol of deuterium, assuming that deuterium is 0.01500% (by number) of natural hydrogen?

35. The power output of the Sun is $4 \times 10^{26} \text{ W}.$

(a) If 90% of this is supplied by the proton-proton cycle, how many protons are consumed per second?

(b) How many neutrinos per second should there be per square meter at the Earth from this process? This huge number is indicative of how rarely a neutrino interacts, since large detectors observe very few per day.

36. Another set of reactions that result in the fusing of hydrogen into helium in the Sun and especially in hotter stars is called the carbon cycle. It is

$$^{12}\text{C} + ^{1}\text{H} \rightarrow ^{13}\text{N} + \gamma,$$

$$^{13}\text{N} \rightarrow ^{13}\text{C} + e^+ + v_e,$$

$$^{13}\text{C} + ^{1}\text{H} \rightarrow ^{14}\text{N} + \gamma,$$

$$^{14}\text{N} + ^{1}\text{H} \rightarrow ^{15}\text{O} + \gamma,$$

$$^{15}\text{O} \rightarrow ^{15}\text{N} + e^+ + v_e.$$

Write down the overall effect of the carbon cycle (as was done for the proton-proton cycle in $2e^- + ^{4}\text{He} \rightarrow ^{4}\text{He} + 2v_e + 6\gamma$). Note the number of protons ($^{1}\text{H}$) required and assume that the positrons ($e^+$) annihilate electrons to form more $\gamma$ rays.

37. (a) Find the total energy released in MeV in each carbon cycle (elaborated in the above problem) including the annihilation energy.

(b) How does this compare with the proton-proton cycle output?

38. Verify that the total number of nucleons, total charge, and electron family number are conserved for each of the fusion reactions in the carbon cycle given in the above problem. (List the value of each of the conserved quantities before and after each of the reactions.)

39. Integrated Concepts

The laser system tested for inertial confinement can produce a 100-kJ pulse only 1.00 ns in duration. (a) What is the power output of the laser system during the brief pulse?

(b) How many photons are in the pulse, given their wavelength is 1.06 $\mu\text{m}?$

(c) What is the total momentum of all these photons?

(d) How does the total photon momentum compare with that of a single 1.00 MeV deuterium nucleus?

40. Integrated Concepts

Find the amount of energy given to the $^{4}\text{He}$ nucleus and to the $\gamma$ ray in the reaction $n + ^{3}\text{He} \rightarrow ^{4}\text{He} + \gamma,$ using the conservation of momentum principle and taking the reactants to be initially at rest. This should confirm the contention that most of the energy goes to the $\gamma$ ray.

41. Integrated Concepts

(a) What temperature gas would have atoms moving fast enough to bring two $^{3}\text{He}$ nuclei into contact? Note that, because both are moving, the average kinetic energy only needs to be half the electric potential energy of these doubly charged nuclei when just in contact with one another.

(b) Does this high temperature imply practical difficulties for doing this in controlled fusion?
42. Integrated Concepts

(a) Estimate the years that the deuterium fuel in the oceans could supply the energy needs of the world. Assume world energy consumption to be ten times that of the United States which is $8 \times 10^{19}$ J/y and that the deuterium in the oceans could be converted to energy with an efficiency of 32%. You must estimate or look up the amount of water in the oceans and take the deuterium content to be 0.015% of natural hydrogen to find the mass of deuterium available. Note that approximate energy yield of deuterium is $3.37 \times 10^{14}$ J/kg.

(b) Comment on how much time this is by any human measure. (It is not an unreasonable result, only an impressive one.)

32.6 Fission

43. (a) Calculate the energy released in the neutron-induced fission (similar to the spontaneous fission in Example 32.3)

$$n + ^{238}\text{U} \rightarrow ^{96}\text{Sr} + ^{140}\text{Xe} + 3n,$$

given $m(^{96}\text{Sr}) = 95.921750$ u and $m(^{140}\text{Xe}) = 139.92164$ u. (b) This result is about 6 MeV greater than the result for spontaneous fission. Why? (c) Confirm that the total number of nucleons and total charge are conserved in this reaction.

44. (a) Calculate the energy released in the neutron-induced fission reaction

$$n + ^{235}\text{U} \rightarrow ^{92}\text{Kr} + ^{142}\text{Ba} + 2n,$$

given $m(^{92}\text{Kr}) = 91.926269$ u and $m(^{142}\text{Ba}) = 141.916361$ u. (b) Confirm that the total number of nucleons and total charge are conserved in this reaction.

45. (a) Calculate the energy released in the neutron-induced fission reaction

$$n + ^{239}\text{Pu} \rightarrow ^{96}\text{Sr} + ^{140}\text{Ba} + 4n,$$

given $m(^{96}\text{Sr}) = 95.921750$ u and $m(^{140}\text{Ba}) = 139.9210581$ u. (b) Confirm that the total number of nucleons and total charge are conserved in this reaction.

46. Confirm that each of the reactions listed for plutonium breeding just following Example 32.4 conserves the total number of nucleons, the total charge, and electron family number.

47. Breeding plutonium produces energy even before any plutonium is fissioned. (The primary purpose of the four nuclear reactors at Chernobyl was breeding plutonium for weapons. Electrical power was a by-product used by the civilian population.) Calculate the energy produced in each of the reactions listed for plutonium breeding just following Example 32.4. The pertinent masses are $m(^{239}\text{U}) = 239.054289$ u, $m(^{235}\text{U}) = 239.054499$ u, $m(^{239}\text{Pu}) = 239.052157$ u, and $m(^{239}\text{Pu}) = 239.052489$ u.

48. The naturally occurring radioactive isotope $^{232}\text{Th}$ does not make good fission fuel, because it has an even number of neutrons; however, it can be bred into a suitable fuel (much as $^{238}\text{U}$ is bred into $^{239}\text{Pu}$). (a) What are $Z$ and $N$ for $^{232}\text{Th}$? (b) Write the reaction equation for neutron captured by $^{232}\text{Th}$ and identify the nuclide $A^XY$ produced in $n + ^{232}\text{Th} \rightarrow A^XY + \gamma$. (c) The product nucleus $\beta^-$ decays, as does its daughter. Write the decay equations for each, and identify the final nucleus. (d) Confirm that the final nucleus has an odd number of neutrons, making it a better fission fuel.

49. The electrical power output of a large nuclear reactor facility is 900 MW. It has a 35.0% efficiency in converting nuclear power to electrical. (a) What is the thermal nuclear power output in megawatts? (b) How many $^{235}\text{U}$ nuclei fission each second, assuming the average fission produces 200 MeV? (c) What mass of $^{235}\text{U}$ is fissioned in one year of full-power operation?

50. A large power reactor that has been in operation for some months is turned off, but residual activity in the core still produces 150 MW of power. If the average energy per decay of the fission products is 1.00 MeV, what is the core activity in curies?

32.7 Nuclear Weapons

51. Find the mass converted into energy by a 12.0-kT bomb.

52. What mass is converted into energy by a 1.00-MT bomb?

53. Fusion bombs use neutrons from their fission trigger to create tritium fuel in the reaction $n + ^{6}\text{Li} \rightarrow ^{3}\text{H} + ^{4}\text{He}$. What is the energy released by this reaction in MeV?

54. It is estimated that the total explosive yield of all the nuclear bombs in existence currently is about 4,000 MT. (a) Convert this amount of energy to kilowatt-hours, noting that 1 kW · h = $3.60 \times 10^7$ J. (b) What would the monetary value of this energy be if it could be converted to electricity costing 10 cents per kW·h?

55. A radiation-enhanced nuclear weapon (or neutron bomb) can have a smaller total yield and still produce more prompt radiation than a conventional nuclear bomb. This allows the use of neutron bombs to kill nearby advancing enemy forces with radiation without blowing up your own forces with the blast. For a 0.500-kT radiation-enhanced weapon and a 1.00-kT conventional nuclear bomb: (a) Compare the blast yields. (b) Compare the prompt radiation yields.

56. (a) How many $^{239}\text{Pu}$ nuclei must fission to produce a 20.0-kT yield, assuming 200 MeV per fission? (b) What is the mass of this much $^{239}\text{Pu}$?

57. Assume one-fourth of the yield of a typical 320-kT strategic bomb comes from fission reactions averaging 200 MeV and the remainder from fusion reactions averaging 20 MeV. (a) Calculate the number of fissions and the approximate mass of uranium and plutonium fissioned, taking the average atomic mass to be 238. (b) Find the number of fusions and calculate the approximate mass of fusion fuel, assuming an average total atomic mass of the two nuclei in each reaction to be 5. (c) Considering the masses found, does it seem reasonable that some missiles could carry 10 warheads? Discuss, noting that the nuclear fuel is only a part of the mass of a warhead.

58. This problem gives some idea of the magnitude of the energy yield of a small tactical bomb. Assume that half the energy of a 1.00-kT nuclear depth charge set off under an aircraft carrier goes into lifting it out of the water—that is, into gravitational potential energy. How high is the carrier lifted if its mass is 90,000 tons?

59. It is estimated that weapons tests in the atmosphere have deposited approximately 9 MCi of $^{90}\text{Sr}$ on the surface of the earth. Find the mass of this amount of $^{90}\text{Sr}$.

60. A 1.00-MT bomb exploded a few kilometers above the ground deposits 25.0% of its energy into radiant heat.
(a) Find the calories per cm$^2$ at a distance of 10.0 km by assuming a uniform distribution over a spherical surface of that radius.

(b) If this heat falls on a person's body, what temperature increase does it cause in the affected tissue, assuming it is absorbed in a layer 1.00-cm deep?

61. Integrated Concepts

One scheme to put nuclear weapons to nonmilitary use is to explode them underground in a geologically stable region and extract the geothermal energy for electricity production. There was a total yield of about 4,000 MT in the combined arsenals in 2006. If 1.00 MT per day could be converted to electricity with an efficiency of 10.0%:

(a) What would the average electrical power output be?

(b) How many years would the arsenal last at this rate?
Learning Objectives

33.1. The Yukawa Particle and the Heisenberg Uncertainty Principle Revisited
- Define Yukawa particle.
- State the Heisenberg uncertainty principle.
- Describe pion.
- Estimate the mass of a pion.
- Explain meson.

33.2. The Four Basic Forces
- State the four basic forces.
- Explain the Feynman diagram for the exchange of a virtual photon between two positive charges.
- Define QED.
- Describe the Feynman diagram for the exchange of a between a proton and a neutron.

33.3. Accelerators Create Matter from Energy
- State the principle of a cyclotron.
- Explain the principle of a synchrotron.
- Describe the voltage needed by an accelerator between accelerating tubes.
- State Fermilab’s accelerator principle.

33.4. Particles, Patterns, and Conservation Laws
- Define matter and antimatter.
- Outline the differences between hadrons and leptons.
- State the differences between mesons and baryons.

33.5. Quarks: Is That All There Is?
- Define fundamental particle.
- Describe quark and antiquark.
- List the flavors of quark.
- Outline the quark composition of hadrons.
- Determine quantum numbers from quark composition.

33.6. GUTs: The Unification of Forces
- State the grand unified theory.
- Explain the electroweak theory.
- Define gluons.
- Describe the principle of quantum chromodynamics.
- Define the standard model.
Introduction to Particle Physics

Following ideas remarkably similar to those of the ancient Greeks, we continue to look for smaller and smaller structures in nature, hoping ultimately to find and understand the most fundamental building blocks that exist. Atomic physics deals with the smallest units of elements and compounds. In its study, we have found a relatively small number of atoms with systematic properties that explained a tremendous range of phenomena. Nuclear physics is concerned with the nuclei of atoms and their substructures. Here, a smaller number of components—the proton and neutron—make up all nuclei. Exploring the systematic behavior of their interactions has revealed even more about matter, force, and energy. Particle physics deals with the substructures of atoms and nuclei and is particularly aimed at finding those truly fundamental particles that have no further substructure. Just as in atomic and nuclear physics, we have found a complex array of particles and properties with systematic characteristics analogous to the periodic table and the chart of nuclides. An underlying structure is apparent, and there is some reason to think that we are finding particles that have no substructure. Of course, we have been in similar situations before. For example, atoms were once thought to be the ultimate substructure. Perhaps we will find deeper and deeper structures and never come to an ultimate substructure. We may never really know, as indicated in Figure 33.2.

This chapter covers the basics of particle physics as we know it today. An amazing convergence of topics is evolving in particle physics. We find that some particles are intimately related to forces, and that nature on the smallest scale may have its greatest influence on the large-scale character of the universe. It is an adventure exceeding the best science fiction because it is not only fantastic, it is real.

### 33.1 The Yukawa Particle and the Heisenberg Uncertainty Principle Revisited

Particle physics as we know it today began with the ideas of Hideki Yukawa in 1935. Physicists had long been concerned with how forces are transmitted, finding the concept of fields, such as electric and magnetic fields to be very useful. A field surrounds an object and carries the force exerted by the object through space. Yukawa was interested in the strong nuclear force in particular and found an ingenious way to explain its short range. His idea is a blend of particles, forces, relativity, and quantum mechanics that is applicable to all forces. Yukawa proposed that force is transmitted by the exchange of particles (called carrier particles). The field consists of these carrier particles.

![Figure 33.2](image_url) The properties of matter are based on substructures called molecules and atoms. Atoms have the substructure of a nucleus with orbiting electrons, the interactions of which explain atomic properties. Protons and neutrons, the interactions of which explain the stability and abundance of elements, form the substructure of nuclei. Protons and neutrons are not fundamental—they are composed of quarks. Like electrons and a few other particles, quarks may be the fundamental building blocks of all there is, lacking any further substructure. But the story is not complete, because quarks and electrons may have substructure smaller than details that are presently observable.

Figure 33.3 illustrates how a pion would carry a force between a proton and a neutron. The pion has mass and can only be created by violating the conservation of mass-energy. This is allowed by the Heisenberg uncertainty principle if it occurs for a sufficiently short period of time. As discussed in Probability: The Heisenberg Uncertainty Principle, the Heisenberg uncertainty principle relates the uncertainties $\Delta E$ in energy and $\Delta t$ in time by

$$\Delta E \Delta t \geq \frac{h}{4\pi},$$

where $h$ is Planck’s constant. Therefore, conservation of mass-energy can be violated by an amount $\Delta E$ for a time $\Delta t \approx \frac{h}{4\pi \Delta E}$ in which time no process can detect the violation. This allows the temporary creation of a particle of mass $m$, where $\Delta E = mc^2$. The larger the mass and the greater the $\Delta E$, the shorter the time it can exist. This means the range of the force is limited, because the particle can only travel a limited distance in a finite amount of time. In fact, the maximum distance is $d \approx c\Delta t$, where $c$ is the speed of light. The pion must then be captured and, thus, cannot be directly observed because that would amount to a permanent violation of mass-energy conservation. Such particles (like the pion above) are called virtual particles, because they cannot be directly observed but their effects can be directly observed. Realizing all this, Yukawa
used the information on the range of the strong nuclear force to estimate the mass of the pion, the particle that carries it. The steps of his reasoning are approximately retraced in the following worked example:

**Example 33.1 Calculating the Mass of a Pion**

Taking the range of the strong nuclear force to be about 1 fermi \((10^{-15} \text{ m})\), calculate the approximate mass of the pion carrying the force, assuming it moves at nearly the speed of light.

**Strategy**

The calculation is approximate because of the assumptions made about the range of the force and the speed of the pion, but also because a more accurate calculation would require the sophisticated mathematics of quantum mechanics. Here, we use the Heisenberg uncertainty principle in the simple form stated above, as developed in *Probability: The Heisenberg Uncertainty Principle*. First, we must calculate the time \(\Delta t\) that the pion exists, given that the distance it travels at nearly the speed of light is about 1 fermi. Then, the Heisenberg uncertainty principle can be solved for the energy \(\Delta E\), and from that the mass of the pion can be determined. We will use the units of \(\text{MeV} / c^2\) for mass, which are convenient since we are often considering converting mass to energy and vice versa.

**Solution**

The distance the pion travels is \(d \approx c\Delta t\), and so the time during which it exists is approximately

\[
\Delta t \approx \frac{d}{c} = \frac{10^{-15} \text{ m}}{3.0 \times 10^8 \text{ m/s}} \approx 3.3 \times 10^{-24} \text{ s}.
\]

Now, solving the Heisenberg uncertainty principle for \(\Delta E\) gives

\[
\Delta E \approx \frac{h}{4\pi\Delta t} \approx \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{4\pi(3.3 \times 10^{-24} \text{ s})} = 100 \text{ MeV}.
\]

Solving this and converting the energy to MeV gives

\[
\Delta E \approx \left(1.6 \times 10^{-11} \right) \frac{1 \text{ MeV}}{1.6 \times 10^{-13} \text{ J}} = 100 \text{ MeV}.
\]

Mass is related to energy by \(\Delta E = mc^2\), so that the mass of the pion is \(m = \Delta E / c^2\), or

\[
m \approx 100 \text{ MeV}/c^2.
\]

**Discussion**

This is about 200 times the mass of an electron and about one-tenth the mass of a nucleon. No such particles were known at the time Yukawa made his bold proposal.

Yukawa’s proposal of particle exchange as the method of force transfer is intriguing. But how can we verify his proposal if we cannot observe the virtual pion directly? If sufficient energy is in a nucleus, it would be possible to free the pion—that is, to create its mass from external energy input. This can be accomplished by collisions of energetic particles with nuclei, but energies greater than 100 MeV are required to conserve both energy and momentum. In 1947, pions were observed in cosmic-ray experiments, which were designed to supply a small flux of high-energy protons that may collide with nuclei. Soon afterward, accelerators of sufficient energy were creating pions in the laboratory under controlled conditions. Three pions were discovered, two with charge and one neutral, and given the symbols \(\pi^+\), \(\pi^-\), and \(\pi^0\), respectively. The masses of \(\pi^+\) and \(\pi^-\) are identical at 139.6 MeV/c^2, whereas \(\pi^0\) has a mass of 135.0 MeV/c^2. These masses are close to the predicted value of 100 MeV/c^2, and since they are intermediate between electron and nucleon masses, the particles are given the name *meson* (now an entire class of particles, as we shall see in *Particles, Patterns, and Conservation Laws*).

The pions, or \(\pi\)-mesons as they are also called, have masses close to those predicted and feel the strong nuclear force. Another previously unknown particle, now called the muon, was discovered during cosmic-ray experiments in 1936 (one of its discoverers, Seth Neddermeyer, also originated the idea of implosion for plutonium bombs). Since the mass of a muon is around 106 MeV/c^2, at first it was thought to be the particle predicted by Yukawa. But it was soon realized that muons do not feel the strong nuclear force and could not be Yukawa’s particle. Their role was unknown, causing the respected physicist I. I. Rabi to comment, “Who ordered that?” This remains a valid question today. We have discovered hundreds of subatomic particles; the roles of some are only partially understood. But there are various patterns and relations to forces that have led to profound insights into nature’s secrets.

### 33.2 The Four Basic Forces

As first discussed in *Problem-Solving Strategies* and mentioned at various points in the text since then, there are only four distinct basic forces in all of nature. This is a remarkably small number considering the myriad phenomena they explain. Particle physics is intimately tied to these four forces. Certain fundamental particles, called carrier particles, carry these forces, and all particles can be classified according to which of the four forces they feel. The table given below summarizes important characteristics of the four basic forces.
<table>
<thead>
<tr>
<th>Force</th>
<th>Approximate relative strength</th>
<th>Range</th>
<th>+/- [1]</th>
<th>Carrier particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravity</td>
<td>$10^{-38}$</td>
<td>$\infty$</td>
<td>+ only</td>
<td>Graviton (conjectured)</td>
</tr>
<tr>
<td>Electromagnetic</td>
<td>$10^{-2}$</td>
<td>$\infty$</td>
<td>+ / -</td>
<td>Photon (observed)</td>
</tr>
<tr>
<td>Weak force</td>
<td>$10^{-13}$</td>
<td>$&lt; 10^{-18}$ m</td>
<td>+ / -</td>
<td>$W^+, W^-, Z^0$ (observed[2])</td>
</tr>
<tr>
<td>Strong force</td>
<td>1</td>
<td>$&lt; 10^{-15}$ m</td>
<td>+ / -</td>
<td>Gluons (conjectured[3])</td>
</tr>
</tbody>
</table>

Figure 33.4: The first image shows the exchange of a virtual photon transmitting the electromagnetic force between charges, just as virtual pion exchange carries the strong nuclear force between nucleons. The second image shows that the photon cannot be directly observed in its passage, because this would disrupt it and alter the force. In this case it does not get to the other charge.

Figure 33.5: The Feynman diagram for the exchange of a virtual photon between two positive charges illustrates how the electromagnetic force is transmitted on a quantum mechanical scale. Time is graphed vertically while the distance is graphed horizontally. The two positive charges are seen to be repelled by the photon exchange.

Although these four forces are distinct and differ greatly from one another under all but the most extreme circumstances, we can see similarities among them. (In GUTs: the Unification of Forces, we will discuss how the four forces may be different manifestations of a single unified force.) Perhaps the most important characteristic among the forces is that they are all transmitted by the exchange of a carrier particle, exactly like what Yukawa had in mind for the strong nuclear force. Each carrier particle is a virtual particle—it cannot be directly observed while transmitting the force.

Figure 33.4 shows the exchange of a virtual photon between two positive charges. The photon cannot be directly observed in its passage, because this would disrupt it and alter the force.

Figure 33.5 shows a way of graphing the exchange of a virtual photon between two positive charges. This graph of time versus position is called a Feynman diagram, after the brilliant American physicist Richard Feynman (1918–1988) who developed it.

Figure 33.6 is a Feynman diagram for the exchange of a virtual pion between a proton and a neutron representing the same interaction as in Figure 33.3. Feynman diagrams are not only a useful tool for visualizing interactions at the quantum mechanical level, they are also used to calculate details of interactions, such as their strengths and probability of occurring. Feynman was one of the theorists who developed the field of quantum electrodynamics (QED), which is the quantum mechanics of electromagnetism. QED has been spectacularly successful in describing electromagnetic interactions on the submicroscopic scale. Feynman was an inspiring teacher, had a colorful personality, and made a profound impact on generations of physicists. He shared the 1965 Nobel Prize with Julian Schwinger and S. i. Tomonaga for work in QED with its deep implications for particle physics.

Why is it that particles called gluons are listed as the carrier particles for the strong nuclear force when, in The Yukawa Particle and the Heisenberg Uncertainty Principle Revisited, we saw that pions apparently carry that force? The answer is that pions are exchanged but they have a

1. + attractive; - repulsive; +/- both.
2. Predicted by theory and first observed in 1983.
Figure 33.6 The image shows a Feynman diagram for the exchange of a $\pi^+$ between a proton and a neutron, carrying the strong nuclear force between them. This diagram represents the situation shown more pictorially in Figure 33.4.

There is a relationship between the mass of the carrier particle and the range of the force. The photon is massless and has energy. So, the existence of (virtual) photons is possible only by virtue of the Heisenberg uncertainty principle and can travel an unlimited distance. Thus, the range of the electromagnetic force is infinite. This is also true for gravity. It is infinite in range because its carrier particle, the graviton, has zero rest mass. (Gravity is the most difficult of the four forces to understand on a quantum scale because it affects the space and time in which the others act. But gravity is so weak that its effects are extremely difficult to observe quantum mechanically. We shall explore it further in General Relativity and Quantum Gravity). The $W^+$, $W^-$, and $Z^0$ particles that carry the weak nuclear force have mass, accounting for the very short range of this force. In fact, the $W^+$, $W^-$, and $Z^0$ are about 1000 times more massive than pions, consistent with the fact that the range of the weak nuclear force is about 1/1000 that of the strong nuclear force. Gluons are actually massless, but since they act inside massive carrier particles like pions, the strong nuclear force is also short ranged.

The relative strengths of the forces given in the Table 33.1 are those for the most common situations. When particles are brought very close together, the relative strengths change, and they may become identical at extremely close range. As we shall see in GUTs: the Unification of Forces, carrier particles may be altered by the energy required to bring particles very close together—in such a manner that they become identical.

### 33.3 Accelerators Create Matter from Energy

Before looking at all the particles we now know about, let us examine some of the machines that created them. The fundamental process in creating previously unknown particles is to accelerate known particles, such as protons or electrons, and direct a beam of them toward a target. Collisions with target nuclei provide a wealth of information, such as information obtained by Rutherford using energetic helium nuclei from natural $\alpha$ radiation. But if the energy of the incoming particles is large enough, new matter is sometimes created in the collision. The more energy input or $\Delta E$, the more matter $m$ can be created, since $m = \Delta E/c^2$. Limitations are placed on what can occur by known conservation laws, such as conservation of mass-energy, momentum, and charge. Even more interesting are the unknown limitations provided by nature. Some expected reactions do occur, while others do not, and still other unexpected reactions may appear. New laws are revealed, and the vast majority of what we know about particle physics has come from accelerator laboratories. It is the particle physicist’s favorite indoor sport, which is partly inspired by theory.

#### Early Accelerators

An early accelerator is a relatively simple, large-scale version of the electron gun. The Van de Graaff (named after the Dutch physicist), which you have likely seen in physics demonstrations, is a small version of the ones used for nuclear research since their invention for that purpose in 1932. For more, see Figure 33.7. These machines are electrostatic, creating potentials as great as 50 MV, and are used to accelerate a variety of nuclei for a range of experiments. Energies produced by Van de Graaffs are insufficient to produce new particles, but they have been instrumental in exploring several aspects of the nucleus. Another, equally famous, early accelerator is the cyclotron, invented in 1930 by the American physicist, E. O. Lawrence (1901–1958). For a visual representation with more detail, see Figure 33.8. Cyclotrons use fixed-frequency alternating electric fields to accelerate particles. The particles spiral outward in a magnetic field, making increasingly larger radius orbits during acceleration. This clever arrangement allows the successive addition of electric potential energy and so greater particle energies are possible than in a Van de Graaff. Lawrence was involved in many early discoveries and in the promotion of physics programs in American universities. He was awarded the 1939 Nobel Prize in Physics for the cyclotron and nuclear activations, and he has an element and two major laboratories named for him.

A synchrotron is a version of a cyclotron in which the frequency of the alternating voltage and the magnetic field strength are increased as the beam particles are accelerated. Particles are made to travel the same distance in a shorter time with each cycle in fixed-radius orbits. A ring of magnets and accelerating tubes, as shown in Figure 33.9, are the major components of synchrotrons. Accelerating voltages are synchronized (i.e., occur at the same time) with the particles to accelerate them, hence the name. Magnetic field strength is increased to keep the orbital radius constant as energy increases. High-energy particles require strong magnetic fields to steer them, so superconducting magnets are commonly employed. Still limited by...
achievable magnetic field strengths, synchrotrons need to be very large at very high energies, since the radius of a high-energy particle’s orbit is very large. Radiation caused by a magnetic field accelerating a charged particle perpendicular to its velocity is called synchrotron radiation in honor of its importance in these machines. Synchrotron radiation has a characteristic spectrum and polarization, and can be recognized in cosmic rays, implying large-scale magnetic fields acting on energetic and charged particles in deep space. Synchrotron radiation produced by accelerators is sometimes used as a source of intense energetic electromagnetic radiation for research purposes.

Figure 33.7 An artist’s rendition of a Van de Graaff generator.

Figure 33.8 Cyclotrons use a magnetic field to cause particles to move in circular orbits. As the particles pass between the plates of the Ds, the voltage across the gap is oscillated to accelerate them twice in each orbit.

Modern Behemoths and Colliding Beams

Physicists have built ever-larger machines, first to reduce the wavelength of the probe and obtain greater detail, then to put greater energy into collisions to create new particles. Each major energy increase brought new information, sometimes producing spectacular progress, motivating the next step. One major innovation was driven by the desire to create more massive particles. Since momentum needs to be conserved in a collision, the particles created by a beam hitting a stationary target should recoil. This means that part of the energy input goes into recoil kinetic energy, significantly limiting the fraction of the beam energy that can be converted into new particles. One solution to this problem is to have head-on collisions between particles moving in opposite directions. Colliding beams are made to meet head-on at points where massive detectors are located. Since the total incoming momentum is zero, it is possible to create particles with momenta and kinetic energies near zero. Particles with masses equivalent to twice the beam energy can thus be created. Another innovation is to create the antimatter counterpart of the beam particle, which thus has the opposite charge and circulates in the opposite direction in the same beam pipe. For a schematic representation, see Figure 33.10.

Figure 33.9 (a) A synchrotron has a ring of magnets and accelerating tubes. The frequency of the accelerating voltages is increased to cause the beam particles to travel the same distance in shorter time. The magnetic field should also be increased to keep each beam burst traveling in a fixed-radius path. Limits on magnetic field strength require these machines to be very large in order to accelerate particles to very high energies. (b) A positive particle is shown in the gap between accelerating tubes. (c) While the particle passes through the tube, the potentials are reversed so that there is another acceleration at the next gap. The frequency of the reversals needs to be varied as the particle is accelerated to achieve successive accelerations in each gap.

This content is available for free at http://cnx.org/content/col11406/1.7
Detectors capable of finding the new particles in the spray of material that emerges from colliding beams are as impressive as the accelerators. While the Fermilab Tevatron had proton and antiproton beam energies of about 1 TeV, so that it can create particles up to $2 \text{ TeV}/c^2$, the Large Hadron Collider (LHC) at the European Center for Nuclear Research (CERN) has achieved beam energies of 3.5 TeV, so that it has a 7-TeV collision energy; CERN hopes to double the beam energy in 2014. The now-canceled Superconducting Super Collider was being constructed in Texas with a design energy of 20 TeV to give a 40-TeV collision energy. It was to be an oval 30 km in diameter. Its cost as well as the politics of international research funding led to its demise.

In addition to the large synchrotrons that produce colliding beams of protons and antiprotons, there are other large electron-positron accelerators. The oldest of these was a straight-line or **linear accelerator**, called the Stanford Linear Accelerator (SLAC), which accelerated particles up to 50 GeV as seen in Figure 33.11. Positrons created by the accelerator were brought to the same energy and collided with electrons in specially designed detectors. Linear accelerators use accelerating tubes similar to those in synchrotrons, but aligned in a straight line. This helps eliminate synchrotron radiation losses, which are particularly severe for electrons made to follow curved paths. CERN had an electron-positron collider appropriately called the Large Electron-Positron Collider (LEP), which accelerated particles to 100 GeV and created a collision energy of 200 GeV. It was 8.5 km in diameter, while the SLAC machine was 3.2 km long.

**Example 33.2 Calculating the Voltage Needed by the Accelerator Between Accelerating Tubes**

A linear accelerator designed to produce a beam of 800-MeV protons has 2000 accelerating tubes. What average voltage must be applied between tubes (such as in the gaps in Figure 33.9) to achieve the desired energy?

**Strategy**

The energy given to the proton in each gap between tubes is $\text{PE}_{\text{elec}} = qV$ where $q$ is the proton’s charge and $V$ is the potential difference (voltage) across the gap. Since $q = q_e = 1.6 \times 10^{-19} \text{ C}$ and $1 \text{ eV} = (1 \text{ V})(1.6 \times 10^{-19} \text{ C})$, the proton gains 1 eV in energy for each volt across the gap that it passes through. The AC voltage applied to the tubes is timed so that it adds to the energy in each gap. The effective voltage is the sum of the gap voltages and equals 800 MV to give each proton an energy of 800 MeV.

**Solution**

There are 2000 gaps and the sum of the voltages across them is 800 MV; thus,

$$V_{\text{gap}} = \frac{800 \text{ MV}}{2000} = 400 \text{ kV}.$$  

$(33.6)$
In the early 1930s only a small number of subatomic particles were known to exist—the proton, neutron, electron, photon and, indirectly, the neutrino. Nature seemed relatively simple in some ways, but mysterious in others. Why, for example, should the particle that carries positive charge be almost 2000 times as massive as the one carrying negative charge? Why does a neutral particle like the neutron have a magnetic moment? Does this imply an internal structure with a distribution of moving charges? Why is it that the electron seems to have no size other than its wavelength, while the proton and neutron are about 1 fermi in size? So, while the number of known particles was small and they explained a great deal of atomic and nuclear phenomena, there were many unexplained phenomena and hints of further substructures.

Things soon became more complicated, both in theory and in the prediction and discovery of new particles. In 1928, the British physicist P.A.M. Dirac (see Figure 33.12) developed a highly successful relativistic quantum theory that laid the foundations of quantum electrodynamics (QED). His theory, for example, explained electron spin and magnetic moment in a natural way. But Dirac’s theory also predicted negative energy states for free electrons. By 1931, Dirac, along with Oppenheimer, realized this was a prediction of positively charged electrons (or positrons). In 1932, American physicist Carl Anderson discovered the positron in cosmic ray studies. The positron, or $e^+$, is the same particle as emitted in $\beta^+$ decay and was the first antimatter that was discovered. In 1935, Yukawa predicted pions as the carriers of the strong nuclear force, and they were eventually discovered. Muons were discovered in cosmic ray experiments in 1937, and they seemed to be heavy, unstable versions of electrons and positrons.

After World War II, accelerators energetic enough to create these particles were built. Not only were predicted and known particles created, but many unexpected particles were observed. Initially called elementary particles, their numbers proliferated to dozens and then hundreds, and the term “particle zoo” became the physicist’s lament at the lack of simplicity. But patterns were observed in the particle zoo that led to simplifying ideas such as quarks, as we shall soon see.

**Figure 33.12** P.A.M. Dirac’s theory of relativistic quantum mechanics not only explained a great deal of what was known, it also predicted antimatter. (credit: Cambridge University, Cavendish Laboratory)

### Matter and Antimatter

The positron was only the first example of antimatter. Every particle in nature has an antimatter counterpart, although some particles, like the photon, are their own antiparticles. Antimatter has charge opposite to that of matter (for example, the positron is positive while the electron is negative) but is nearly identical otherwise, having the same mass, intrinsic spin, half-life, and so on. When a particle and its antimatter counterpart interact, they annihilate one another, usually totally converting their masses to pure energy in the form of photons as seen in Figure 33.13. Neutral particles, such as neutrons, have neutral antimatter counterparts, which also annihilate when they interact. Certain neutral particles are their own antiparticle and live correspondingly short lives. For example, the neutral pion $\pi^0$ is its own antiparticle and has a half-life about $10^{-8}$ shorter than $\pi^+$ and $\pi^-$, which are each other’s antiparticles. Without exception, nature is symmetric—all particles have antimatter counterparts. For example, antiprotons and antineutrons were first created in accelerator experiments in 1956 and the antiproton is negative. Antihydrogen atoms, consisting of an antiproton and antineutron, were observed in 1995 at CERN, too. It is possible to contain large-scale antimatter particles such as antiprotons by using electromagnetic traps that confine the particles within a magnetic field so that they don’t annihilate with other particles. However, particles of the same charge repel each other, so the more particles that are contained in a trap, the more energy is needed to power the magnetic field that contains them. It is not currently possible to store a significant quantity of antiprotons. At any rate, we now see that negative charge is associated with both low-mass (electrons) and high-mass particles (antiprotons) and the apparent asymmetry is not there. But this knowledge does raise another question—why is there such a predominance of matter and so little antimatter? Possible explanations emerge later in this and the next chapter.

### Hadrons and Leptons

Particles can also be revealingly grouped according to what forces they feel between them. All particles (even those that are massless) are affected by gravity, since gravity affects the space and time in which particles exist. All charged particles are affected by the electromagnetic force, as are neutral particles that have an internal distribution of charge (such as the neutron with its magnetic moment). Special names are given to particles that feel the strong and weak nuclear forces. **Hadrons** are particles that feel the strong nuclear force, whereas **leptons** are particles that do not.
proton, neutron, and the pions are examples of hadrons. The electron, positron, muons, and neutrinos are examples of leptons, the name meaning low mass. Leptons feel the weak nuclear force. In fact, all particles feel the weak nuclear force. This means that hadrons are distinguished by being able to feel both the strong and weak nuclear forces.

Table 33.2 lists the characteristics of some of the most important subatomic particles, including the directly observed carrier particles for the electromagnetic and weak nuclear forces, all leptons, and some hadrons. Several hints related to an underlying substructure emerge from an examination of these particle characteristics. Note that the carrier particles are called **gauge bosons**. First mentioned in Patterns in Spectra Reveal More Quantization, a **boson** is a particle with zero or an integer value of intrinsic spin (such as \( s = 0, 1, 2, ... \)), whereas a **fermion** is a particle with a half-integer value of intrinsic spin (\( s = 1/2, 3/2, ... \)). Fermions obey the Pauli exclusion principle whereas bosons do not. All the known and conjectured carrier particles are bosons.

Figure 33.13 When a particle encounters its antiparticle, they annihilate, often producing pure energy in the form of photons. In this case, an electron and a positron convert all their mass into two identical energy rays, which move away in opposite directions to keep total momentum zero as it was before. Similar annihilations occur for other combinations of a particle with its antiparticle, sometimes producing more particles while obeying all conservation laws.
### Table 33.2 Selected Particle Characteristics

<table>
<thead>
<tr>
<th>Category</th>
<th>Particle name</th>
<th>Symbol</th>
<th>Antiparticle</th>
<th>Rest mass (MeV/c²)</th>
<th>B</th>
<th>L_e</th>
<th>L_µ</th>
<th>L_τ</th>
<th>S</th>
<th>Lifetime[^5] (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gauge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stable</td>
</tr>
<tr>
<td><strong>Bosons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>W⁺</td>
<td>W⁻</td>
<td>80.39×10³</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.6×10⁻²⁵</td>
</tr>
<tr>
<td>Z</td>
<td>Z⁰</td>
<td>Self</td>
<td>91.19×10³</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.32×10⁻²⁵</td>
</tr>
<tr>
<td><strong>Leptons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electron</td>
<td>e⁻</td>
<td>e⁺</td>
<td>0.511</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Stable</td>
</tr>
<tr>
<td>Neutrino (e)</td>
<td>νₑ</td>
<td>νₑ</td>
<td>0(7.0eV)</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Stable</td>
</tr>
<tr>
<td>Muon</td>
<td>μ⁻</td>
<td>μ⁺</td>
<td>105.7</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.20×10⁻⁶</td>
</tr>
<tr>
<td>Neutrino (μ)</td>
<td>ν_μ</td>
<td>ν_μ</td>
<td>0( &lt; 0.27)</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Stable</td>
</tr>
<tr>
<td>Tau</td>
<td>τ⁻</td>
<td>τ⁺</td>
<td>1777</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>2.91×10⁻¹³</td>
</tr>
<tr>
<td>Neutrino (τ)</td>
<td>ν_τ</td>
<td>ν_τ</td>
<td>0( &lt; 31)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>Stable</td>
</tr>
<tr>
<td><strong>Hadrons (selected)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mesons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pion</td>
<td>π⁺</td>
<td>π⁻</td>
<td>139.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.60×10⁻⁸</td>
</tr>
<tr>
<td></td>
<td>π⁰</td>
<td>Self</td>
<td>135.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8.4×10⁻¹⁷</td>
</tr>
<tr>
<td>Kaon</td>
<td>K⁺</td>
<td>K⁻</td>
<td>493.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>1.24×10⁻⁸</td>
</tr>
<tr>
<td></td>
<td>K⁰</td>
<td>K⁰⁻</td>
<td>497.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0.90×10⁻¹⁰</td>
</tr>
<tr>
<td>Eta</td>
<td>η⁰</td>
<td>Self</td>
<td>547.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.53×10⁻¹⁹</td>
</tr>
<tr>
<td></td>
<td>(many other mesons known)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baryons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proton</td>
<td>p</td>
<td>p⁻</td>
<td>938.3</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Stable[^7]</td>
</tr>
<tr>
<td>Neutron</td>
<td>n</td>
<td>n⁻</td>
<td>939.6</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>882</td>
</tr>
<tr>
<td>Lambda</td>
<td>Λ⁰</td>
<td>Λ⁻</td>
<td>1115.7</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>2.63×10⁻¹⁰</td>
</tr>
<tr>
<td>Sigma</td>
<td>Σ⁺</td>
<td>Σ⁻</td>
<td>1189.4</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>0.80×10⁻¹⁰</td>
</tr>
<tr>
<td></td>
<td>Σ⁰</td>
<td>Σ⁰⁻</td>
<td>1192.6</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>7.4×10⁻¹⁰⁻²⁰</td>
</tr>
<tr>
<td></td>
<td>Σ⁻</td>
<td>Σ⁺</td>
<td>1197.4</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 1</td>
<td>0</td>
<td>1.48×10⁻¹⁰</td>
</tr>
<tr>
<td>Xi</td>
<td>Ξ⁰</td>
<td>Ξ⁻</td>
<td>1314.9</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 2</td>
<td>0</td>
<td>2.90×10⁻¹⁰</td>
</tr>
<tr>
<td></td>
<td>Ξ⁻</td>
<td>Ξ⁺</td>
<td>1321.7</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 2</td>
<td>0</td>
<td>1.64×10⁻¹⁰</td>
</tr>
<tr>
<td>Omega</td>
<td>Ω⁻</td>
<td>Ω⁺</td>
<td>1672.5</td>
<td>± 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>± 3</td>
<td>0</td>
<td>0.82×10⁻¹⁰</td>
</tr>
<tr>
<td></td>
<td>(many other baryons known)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All known leptons are listed in the table given above. There are only six leptons (and their antiparticles), and they seem to be fundamental in that they have no apparent underlying structure. Leptons have no discernible size other than their wavelength, so that we know they are pointlike down to about 10⁻¹⁸ m. The leptons fall into three families, implying three conservation laws for three quantum numbers. One of these was known from β decay, where the existence of the electron’s neutrino implied that a new quantum number, called the electron family number Lₑ is conserved. Thus, in β decay, an antielectron’s neutrino \(\bar{\nu}_e\) must be created with \(L_e = -1\) when an electron with \(L_e = +1\) is created, so that the total remains 0 as it was before decay.

[^4]: The lower of the \(\mp\) or ± symbols are the values for antiparticles.
[^5]: Lifetimes are traditionally given as \(t_{1/2}/0.693\) (which is \(1/\lambda\), the inverse of the decay constant).
[^6]: Neutrino masses may be zero. Experimental upper limits are given in parentheses.
[^7]: Experimental lower limit is \(>5\times10^{-32}\) for proposed mode of decay.
Once the muon was discovered in cosmic rays, its decay mode was found to be

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu,$$

which implied another “family” and associated conservation principle. The particle $\nu_\mu$ is a muon’s neutrino, and it is created to conserve muon family number $L_\mu$. So muons are leptons with a family of their own, and conservation of total $L_\mu$ also seems to be obeyed in many experiments.

More recently, a third lepton family was discovered when $\tau$ particles were created and observed to decay in a manner similar to muons. One principal decay mode is

$$\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau.$$  

(33.8)

Conservation of total $L_\tau$ seems to be another law obeyed in many experiments. In fact, particle experiments have found that lepton family number is not universally conserved, due to neutrino “oscillations,” or transformations of neutrinos from one family type to another.

Mesons and Baryons

Now, note that the hadrons in the table given above are divided into two subgroups, called mesons (originally for medium mass) and baryons (the name originally meaning large mass). The division between mesons and baryons is actually based on their observed decay modes and is not strictly associated with their masses. Mesons are hadrons that can decay to leptons and leave no hadrons, which implies that mesons are not conserved in number. Baryons are hadrons that always decay to another baryon. A new physical quantity called baryon number $B$ seems to always be conserved in nature and is listed for the various particles in the table above. Mesons and leptons have $B = 0$ so that they can decay to other particles with $B = 0$. But baryons have $B = +1$ if they are matter, and $B = -1$ if they are antimatter. The conservation of total baryon number is a more general rule than first noted in nuclear physics, where it was observed that the total number of nucleons was always conserved in nuclear reactions and decays. That rule in nuclear physics is just one consequence of the conservation of the total baryon number.

Forces, Reactions, and Reaction Rates

The forces that act between particles regulate how they interact with other particles. For example, pions feel the strong force and do not penetrate as far in matter as do muons, which do not feel the strong force. (This was the way those who discovered the muon knew it could not be the particle that carries the strong force—its penetration or range was too great for it to be feeling the strong force.) Similarly, reactions that create other particles, like cosmic rays interacting with nuclei in the atmosphere, have greater probability if they are caused by the strong force than if they are caused by the weak force. Such knowledge has been useful to physicists while analyzing the particles produced by various accelerators.

The forces experienced by particles also govern how particles interact with themselves if they are unstable and decay. For example, the stronger the force, the faster they decay and the shorter is their lifetime. An example of a nuclear decay via the strong force is $^8\text{Be} \rightarrow \alpha + \alpha$ with a lifetime of about $10^{-16}$ s. The neutron is a good example of decay via the weak force. The process $n \rightarrow p + e^- + \bar{\nu}_e$ has a longer lifetime of 882 s. The weak force causes this decay, as it does all $\beta$ decay. An important clue that the weak force is responsible for $\beta$ decay is the creation of leptons, such as $e^-$ and $\bar{\nu}_e$. None would be created if the strong force was responsible, just as no leptons are created in the decay of $^8\text{Be}$. The systematics of particle lifetimes is a little simpler than nuclear lifetimes when hundreds of particles are examined (not just the ones in the table given above). Particles that decay via the weak force have lifetimes mostly in the range of $10^{-16}$ to $10^{-12}$ s, whereas those that decay via the strong force have lifetimes mostly in the range of $10^{-16}$ to $10^{-23}$ s. Turning this around, if we measure the lifetime of a particle, we can tell if it decays via the weak or strong force.

Yet another quantum number emerges from decay lifetimes and patterns. Note that the particles $\Lambda, \Sigma, \Xi$, and $\Omega$ decay with lifetimes on the order of $10^{-10}$ s (the exception is $\Sigma^0$, whose short lifetime is explained by its particular quark substructure), implying that their decay is caused by the weak force alone, although they are hadrons and feel the strong force. The decay modes of these particles also show patterns—in particular, certain decays that should be possible within all the known conservation laws do not occur. Whenever something is possible in physics, it will happen. If something does not happen, it is forbidden by a rule. All this seemed strange to those studying these particles when they were first discovered, so they named a new quantum number strangeness, given the symbol $S$ in the table given above. The values of strangeness assigned to various particles are based on the decay systematics. It is found that strangeness is conserved by the strong force, which governs the production of most of these particles in accelerator experiments. However, strangeness is not conserved by the weak force. This conclusion is reached from the fact that particles that have long lifetimes decay via the weak force and do not conserve strangeness. All of this also has implications for the carrier particles, since they transmit forces and are thus involved in these decays.

**Example 33.3 Calculating Quantum Numbers in Two Decays**

(a) The most common decay mode of the $\Xi^-$ particle is $\Xi^- \rightarrow \Lambda^0 + \pi^-$. Using the quantum numbers in the table given above, show that strangeness changes by 1, baryon number and charge are conserved, and lepton family numbers are unaffected.

(b) Is the decay $K^+ \rightarrow \mu^+ + \nu_\mu$ allowed, given the quantum numbers in the table given above?

**Strategy**

In part (a), the conservation laws can be examined by adding the quantum numbers of the decay products and comparing them with the parent particle. In part (b), the same procedure can reveal if a conservation law is broken or not.
Solution for (a)
Before the decay, the $\Xi^-$ has strangeness $S = -2$. After the decay, the total strangeness is $-1$ for the $\Lambda^0$, plus $0$ for the $\pi^-$. Thus, total strangeness has gone from $-2$ to $-1$ or a change of $+1$. Baryon number for the $\Xi^-$ is $B = +1$ before the decay, and after the decay the $\Lambda^0$ has $B = +1$ and the $\pi^-$ has $B = 0$ so that the total baryon number remains $+1$. Charge is $-1$ before the decay, and the total charge after is also $0 - 1 = -1$. Lepton numbers for all the particles are zero, and so lepton numbers are conserved.

Discussion for (a)
The $\Xi^-$ decay is caused by the weak interaction, since strangeness changes, and it is consistent with the relatively long $1.64 \times 10^{-10}$-s lifetime of the $\Xi^-$. 

Solution for (b)
The decay $K^+ \rightarrow \mu^+ + \nu_\mu$ is allowed if charge, baryon number, mass-energy, and lepton numbers are conserved. Strangeness can change due to the weak interaction. Charge is conserved as $s \rightarrow d$. Baryon number is conserved, since all particles have $B = 0$. Mass-energy is conserved in the sense that the $K^+$ has a greater mass than the products, so that the decay can be spontaneous. Lepton family numbers are conserved at 0 for the electron and tau family for all particles. The muon family number is $L_\mu = 0$ before and $L_\mu = -1 + 1 = 0$ after. Strangeness changes from $+1$ before to $0 + 0$ after, for an allowed change of 1. The decay is allowed by all these measures.

Discussion for (b)
This decay is not only allowed by our reckoning, it is, in fact, the primary decay mode of the $K^+$ meson and is caused by the weak force, consistent with the long $1.24 \times 10^{-8}$-s lifetime.

There are hundreds of particles, all hadrons, not listed in Table 33.2, most of which have shorter lifetimes. The systematics of those particle lifetimes, their production probabilities, and decay products are completely consistent with the conservation laws noted for lepton families, baryon number, and strangeness, but they also imply other quantum numbers and conservation laws. There are a finite, and in fact relatively small, number of these conserved quantities, however, implying a finite set of substructures. Additionally, some of these short-lived particles resemble the excited states of other particles, implying an internal structure. All of this jigsaw puzzle can be tied together and explained relatively simply by the existence of fundamental substructures. Leptons seem to be fundamental structures. Hadrons seem to have a substructure called quarks. Quarks: Is That All There Is? explores the basics of the underlying quark building blocks.

Figure 33.14 Murray Gell-Mann (b. 1929) proposed quarks as a substructure of hadrons in 1963 and was already known for his work on the concept of strangeness. Although quarks have never been directly observed, several predictions of the quark model were quickly confirmed, and their properties explain all known hadron characteristics. Gell-Mann was awarded the Nobel Prize in 1969. (credit: Luboš Motl)

33.5 Quarks: Is That All There Is?
Quarks have been mentioned at various points in this text as fundamental building blocks and members of the exclusive club of truly elementary particles. Note that an elementary or fundamental particle has no substructure (it is not made of other particles) and has no finite size other than its wavelength. This does not mean that fundamental particles are stable—some decay, while others do not. Keep in mind that all leptons seem to be
fundamental, whereas no hadrons are fundamental. There is strong evidence that quarks are the fundamental building blocks of hadrons as seen in Figure 33.15. Quarks are the second group of fundamental particles (leptons are the first). The third and perhaps final group of fundamental particles is the carrier particles for the four basic forces. Leptons, quarks, and carrier particles may be all there is. In this module we will discuss the quark substructure of hadrons and its relationship to forces as well as indicate some remaining questions and problems.

![Figure 33.15](image_url)

**Figure 33.15** All baryons, such as the proton and neutron shown here, are composed of three quarks. All mesons, such as the pions shown here, are composed of a quark-antiquark pair. Arrows represent the spins of the quarks, which, as we shall see, are also colored. The colors are such that they need to add to white for any possible combination of quarks.

**Conception of Quarks**

Quarks were first proposed independently by American physicists Murray Gell-Mann and George Zweig in 1963. Their quaint name was taken by Gell-Mann from a James Joyce novel—Gell-Mann was also largely responsible for the concept and name of strangeness. (Whimsical names are common in particle physics, reflecting the personalities of modern physicists.) Originally, three quark types—or flavors—were proposed to account for the then-known mesons and baryons. These quark flavors are named up ($u$), down ($d$), and strange ($s$). All quarks have half-integral spin and are thus fermions. All mesons have integral spin while all baryons have half-integral spin. Therefore, mesons should be made up of an even number of quarks while baryons need to be made up of an odd number of quarks. Figure 33.15 shows the quark substructure of the proton, neutron, and two pions. The most radical proposal by Gell-Mann and Zweig is the fractional charges of quarks, which are $\pm \frac{2}{3} q_e$ and $\pm \frac{1}{3} q_e$, whereas all directly observed particles have charges that are integral multiples of $q_e$. Note that the fractional value of the quark does not violate the fact that the $e$ is the smallest unit of charge that is observed, because a free quark cannot exist. Table 33.3 lists characteristics of the six quark flavors that are now thought to exist. Discoveries made since 1963 have required extra quark flavors, which are divided into three families quite analogous to leptons.

**How Does it Work?**

To understand how these quark substructures work, let us specifically examine the proton, neutron, and the two pions pictured in Figure 33.15 before moving on to more general considerations. First, the proton $p$ is composed of the three quarks $uud$, so that its total charge is $+\left(\frac{2}{3}\right) q_e + +\left(\frac{2}{3}\right) q_e - -\left(\frac{1}{3}\right) q_e = q_e$, as expected. With the spins aligned as in the figure, the proton’s intrinsic spin is $+\left(\frac{1}{2}\right) + +\left(\frac{1}{2}\right) - -\left(\frac{1}{2}\right) = +\left(\frac{3}{2}\right)$, also as expected. Note that the spins of the up quarks are aligned, so that they would be in the same state except that they have different colors (another quantum number to be elaborated upon a little later). Quarks obey the Pauli exclusion principle. Similar comments apply to the neutron $n$, which is composed of the three quarks $udd$. Note also that the neutron is made of charges that add to zero but move internally, producing its well-known magnetic moment. When the neutron $\beta^-$ decays, it does so by changing the flavor of one of its quarks. Writing neutron $\beta^-$ decay in terms of quarks,

$$n \rightarrow p + \beta^- + \bar{\nu}_e \text{ becomes } uud \rightarrow uud + \beta^- + \bar{\nu}_e.$$

We see that this is equivalent to a down quark changing flavor to become an up quark:

$$d \rightarrow u + \beta^- + \bar{\nu}_e.$$

(33.9)
<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Antiparticle</th>
<th>Spin</th>
<th>Charge</th>
<th>B</th>
<th>S</th>
<th>c</th>
<th>b</th>
<th>t</th>
<th>Mass (GeV/c²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>u</td>
<td>ū</td>
<td>1/2</td>
<td>±2q/e</td>
<td>±1/3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.005</td>
</tr>
<tr>
<td>Down</td>
<td>d</td>
<td>ṗ</td>
<td>1/2</td>
<td>+1q/e</td>
<td>±1/3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.008</td>
</tr>
<tr>
<td>Strange</td>
<td>s</td>
<td>̅s</td>
<td>1/2</td>
<td>+1q/e</td>
<td>±1/3</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.50</td>
</tr>
<tr>
<td>Charmed</td>
<td>c</td>
<td>̅c</td>
<td>1/2</td>
<td>±2q/e</td>
<td>±1/3</td>
<td>0</td>
<td>±1</td>
<td>0</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Bottom</td>
<td>b</td>
<td>̅b</td>
<td>1/2</td>
<td>+1q/e</td>
<td>±1/3</td>
<td>0</td>
<td>0</td>
<td>±1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Top</td>
<td>t</td>
<td>̅t</td>
<td>1/2</td>
<td>±2q/e</td>
<td>±1/3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>±1</td>
<td>173</td>
</tr>
</tbody>
</table>

8. The lower of the ± symbols are the values for antiquarks.
9. B is baryon number, S is strangeness, c is charm, b is bottomness, t is topness.
10. Values are approximate, are not directly observable, and vary with model.
### Table 33.4 Quark Composition of Selected Hadrons\[11\]

<table>
<thead>
<tr>
<th>Particle</th>
<th>Quark Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mesons</strong></td>
<td></td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>$u \bar{d}$</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>$\bar{u} d$</td>
</tr>
<tr>
<td>$\pi^0$</td>
<td>$u \bar{u}, , d \bar{d}$ mixture[12]</td>
</tr>
<tr>
<td>$\eta^0$</td>
<td>$u \bar{u}, , d \bar{d}$ mixture[13]</td>
</tr>
<tr>
<td>$K^0$</td>
<td>$d \bar{s}$</td>
</tr>
<tr>
<td>$\bar{K}^0$</td>
<td>$\bar{d} s$</td>
</tr>
<tr>
<td>$K^+$</td>
<td>$u \bar{s}$</td>
</tr>
<tr>
<td>$K^-$</td>
<td>$\bar{u} s$</td>
</tr>
<tr>
<td>$J/\psi$</td>
<td>$c \bar{c}$</td>
</tr>
<tr>
<td>$\Upsilon$</td>
<td>$b \bar{b}$</td>
</tr>
<tr>
<td><strong>Baryons</strong></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>$u u d$</td>
</tr>
<tr>
<td>$n$</td>
<td>$d u d$</td>
</tr>
<tr>
<td>$\Delta^0$</td>
<td>$u d d$</td>
</tr>
<tr>
<td>$\Delta^+$</td>
<td>$u u d$</td>
</tr>
<tr>
<td>$\Delta^-$</td>
<td>$d d d$</td>
</tr>
<tr>
<td>$\Delta^{++}$</td>
<td>$u u u$</td>
</tr>
<tr>
<td>$\Lambda^0$</td>
<td>$u d s$</td>
</tr>
<tr>
<td>$\Sigma^0$</td>
<td>$u d s$</td>
</tr>
<tr>
<td>$\Sigma^+$</td>
<td>$u u s$</td>
</tr>
<tr>
<td>$\Sigma^-$</td>
<td>$d d s$</td>
</tr>
<tr>
<td>$\Xi^0$</td>
<td>$u s s$</td>
</tr>
<tr>
<td>$\Xi^-$</td>
<td>$d s s$</td>
</tr>
<tr>
<td>$\Omega^-$</td>
<td>$s s s$</td>
</tr>
</tbody>
</table>

This is an example of the general fact that the weak nuclear force can change the flavor of a quark. By general, we mean that any quark can be converted to any other (change flavor) by the weak nuclear force. Not only can we get $d \to u$, we can also get $u \to d$. Furthermore, the strange quark can be changed by the weak force, too, making $s \to u$ and $s \to d$ possible. This explains the violation of the conservation of strangeness by the weak force noted in the preceding section. Another general fact is that the strong nuclear force cannot change the flavor of a quark.

Again, from Figure 33.15, we see that the $\pi^+$ meson (one of the three pions) is composed of an up quark plus an antidown quark, or $u \bar{d}$. Its total charge is thus $+\left(\frac{2}{3}\right)q_e + \left(\frac{1}{3}\right)q_e = q_e$, as expected. Its baryon number is 0, since it has a quark and an antiquark with baryon numbers $+\left(\frac{1}{3}\right) - \left(\frac{1}{3}\right) = 0$. The $\pi^+$ half-life is relatively long since, although it is composed of matter and antimatter, the quarks are different flavors and the

---

11. These two mesons are different mixtures, but each is its own antiparticle, as indicated by its quark composition.
12. These two mesons are different mixtures, but each is its own antiparticle, as indicated by its quark composition.
13. These two mesons are different mixtures, but each is its own antiparticle, as indicated by its quark composition.
14. Antibaryons have the antiquarks of their counterparts. The antiproton $\bar{p}$ is $\bar{u} \bar{u} \bar{d}$, for example.
15. Baryons composed of the same quarks are different states of the same particle. For example, the $\Delta^+$ is an excited state of the proton.
weak force should cause the decay by changing the flavor of one into that of the other. The spins of the \( u \) and \( d \) quarks are antiparallel, enabling the pion to have spin zero, as observed experimentally. Finally, the \( \pi^- \) meson shown in Figure 33.15 is the antiparticle of the \( \pi^+ \) meson, and it is composed of the corresponding quark antiparticles. That is, the \( \pi^+ \) meson is \( ud \), while the \( \pi^- \) meson is \( \bar{u}\bar{d} \). These two pions annihilate each other quickly, because their constituent quarks are each other’s antiparticles.

Two general rules for combining quarks to form hadrons are:

1. Baryons are composed of three quarks, and antibaryons are composed of three antiquarks.
2. Mesons are combinations of a quark and an antiquark.

One of the clever things about this scheme is that only integral charges result, even though the quarks have fractional charge.

**All Combinations are Possible**

All quark combinations are possible. Table 33.4 lists some of these combinations. When Gell-Mann and Zweig proposed the original three quark flavors, particles corresponding to all combinations of those three had not been observed. The pattern was there, but it was incomplete—much as had been the case in the periodic table of the elements and the chart of nuclides. The \( \Omega^- \) particle, in particular, had not been discovered but was predicted by quark theory. Its combination of three strange quarks, \( sss \), gives it a strangeness of \( -3 \) (see Table 33.2) and other predictable characteristics, such as spin, charge, approximate mass, and lifetime. If the quark picture is complete, the \( \Omega^- \) should exist. It was first observed in 1964 at Brookhaven National Laboratory and had the predicted characteristics as seen in Figure 33.16. The discovery of the \( \Omega^- \) was convincing indirect evidence for the existence of the three original quark flavors and boosted theoretical and experimental efforts to further explore particle physics in terms of quarks.

**Patterns and Puzzles: Atoms, Nuclei, and Quarks**

Patterns in the properties of atoms allowed the periodic table to be developed. From it, previously unknown elements were predicted and observed. Similarly, patterns were observed in the properties of nuclei, leading to the chart of nuclides and successful predictions of previously unknown nuclides. Now with particle physics, patterns imply a quark substructure that, if taken literally, predicts previously unknown particles. These have now been observed in another triumph of underlying unity.

**Figure 33.16** The image relates to the discovery of the \( \Omega^- \). It is a secondary reaction in which an accelerator-produced \( K^- \) collides with a proton via the strong force and conserves strangeness to produce the \( \Omega^- \) with characteristics predicted by the quark model. As with other predictions of previously unobserved particles, this gave a tremendous boost to quark theory. (credit: Brookhaven National Laboratory)

**Example 33.4 Quantum Numbers From Quark Composition**

Verify the quantum numbers given for the \( \Xi^0 \) particle in Table 33.2 by adding the quantum numbers for its quark composition as given in Table 33.4.

**Strategy**
The composition of the $\Xi^0$ is given as $uss$ in Table 33.4. The quantum numbers for the constituent quarks are given in Table 33.3. We will not consider spin, because that is not given for the $\Xi^0$. But we can check on charge and the other quantum numbers given for the quarks.

**Solution**

The total charge of $uss$ is $+\left(\frac{2}{3}\right)q_e - \left(\frac{1}{3}\right)q_e - \left(\frac{1}{3}\right)q_e = 0$, which is correct for the $\Xi^0$. The baryon number is $+\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right) + \left(\frac{1}{3}\right) = 1$, also correct since the $\Xi^0$ is a matter baryon and has $B = 1$, as listed in Table 33.2. Its strangeness is $S = 0 - 1 - 1 = -2$, also as expected from Table 33.2. Its charm, bottomness, and topness are 0, as are its lepton family numbers (it is not a lepton).

**Discussion**

This procedure is similar to what the inventors of the quark hypothesis did when checking to see if their solution to the puzzle of particle patterns was correct. They also checked to see if all combinations were known, thereby predicting the previously unobserved $\Omega^-$ as the completion of a pattern.

---

**Now, Let Us Talk About Direct Evidence**

At first, physicists expected that, with sufficient energy, we should be able to free quarks and observe them directly. This has not proved possible. There is no direct observation of a fractional charge or any isolated quark. When large energies are put into collisions, other particles are created—but no quarks emerge. There is nearly direct evidence for quarks that is quite compelling. By 1967, experiments at SLAC scattering 20-GeV electrons from protons had produced results like Rutherford had obtained for the nucleus nearly 60 years earlier. The SLAC scattering experiments showed unambiguously that there were three pointlike (meaning they had sizes considerably smaller than the probe’s wavelength) charges inside the proton as seen in Figure 33.17. This evidence made all but the most skeptical admit that there was validity to the quark substructure of hadrons.

---

**Figure 33.17** Scattering of high-energy electrons from protons at facilities like SLAC produces evidence of three point-like charges consistent with proposed quark properties. This experiment is analogous to Rutherford’s discovery of the small size of the nucleus by scattering $\alpha$ particles. High-energy electrons are used so that the probe wavelength is small enough to see details smaller than the proton.

More recent and higher-energy experiments have produced jets of particles in collisions, highly suggestive of three quarks in a nucleon. Since the quarks are very tightly bound, energy put into separating them pulls them only so far apart before it starts being converted into other particles. More energy produces more particles, not a separation of quarks. Conservation of momentum requires that the particles come out in jets along the three paths in which the quarks were being pulled. Note that there are only three jets, and that other characteristics of the particles are consistent with the three-quark substructure.

---

**Figure 33.18** Simulation of a proton-proton collision at 14-TeV center-of-mass energy in the ALICE detector at CERN LHC. The lines follow particle trajectories and the cyan dots represent the energy depositions in the sensitive detector elements. (credit: Matevž Tadel)
Quarks Have Their Ups and Downs

The quark model actually lost some of its early popularity because the original model with three quarks had to be modified. The up and down quarks seemed to compose normal matter as seen in Table 33.4, while the single strange quark explained strangeness. Why didn’t it have a counterpart? A fourth quark flavor called charm \((c)\) was proposed as the counterpart of the strange quark to make things symmetric—there would be two normal quarks \((u \text{ and } d)\) and two exotic quarks \((s \text{ and } c)\). Furthermore, at that time only four leptons were known, two normal and two exotic. It was attractive that there would be four quarks and four leptons. The problem was that no known particles contained a charmed quark. Suddenly, in November of 1974, two groups (one headed by C. C. Ting at Brookhaven National Laboratory and the other by Burton Richter at SLAC) independently and nearly simultaneously discovered a new meson with characteristics that made it clear that its substructure is \(c \bar{c}\). It was called \(J/\psi\) by one group and psi \((\psi)\) by the other and now is known as the \(J/\psi\) meson. Since then, numerous particles have been discovered containing the charmed quark, consistent in every way with the quark model. The discovery of the \(J/\psi\) meson had such a rejuvenating effect on quark theory that it is now called the November Revolution. Ting and Richter shared the 1976 Nobel Prize.

History quickly repeated itself. In 1975, the tau \((\tau)\) was discovered, and a third family of leptons emerged as seen in Table 33.2. Theorists quickly proposed two more quark flavors called top \((t)\) or truth and bottom \((b)\) or beauty to keep the number of quarks the same as the number of leptons. And in 1976, the upsilon \((\Upsilon)\) meson was discovered and shown to be composed of a bottom and an antitop quark or \(b \bar{b}\), quite analogous to the \(J/\psi\) being \(c \bar{c}\) as seen in Table 33.4. Being a single meson, these sons are sometimes called bare charm and bare bottom and reveal the characteristics of their quarks most clearly. Other mesons containing bottom quarks have since been observed. In 1995, two groups at Fermilab confirmed the top quark’s existence, completing the picture of six quarks listed in Table 33.3. Each successive quark discovery—first \(c\), then \(b\), and finally \(t\)—has required higher energy because each has higher mass. Quark masses in Table 33.3 are only approximately known, because they are not directly observed. They must be inferred from the masses of the particles they combine to form.

What’s Color got to do with it?—A Whiter Shade of Pale

As mentioned and shown in Figure 33.15, quarks carry another quantum number, which we call color. Of course, it is not the color we sense with visible light, but its properties are analogous to those of three primary and three secondary colors. Specifically, a quark can have one of three color values we call red \((R)\), green \((G)\), and blue \((B)\) in analogy to those primary visible colors. Antiquarks have three values we call antired or cyan \((\bar{R})\), antigreen or magenta \((\bar{G})\), and antiblue or yellow \((\bar{B})\) in analogy to those secondary visible colors. The reason for these names is that when certain visual colors are combined, the eye sees white. The analogy of the colors combining to white is used to explain why baryons are made of three quarks, why mesons are a quark and an antiquark, and why we cannot isolate a single quark. The force between the quarks is such that their combined colors produce white. This is illustrated in Figure 33.19. A baryon must have one of each primary color or RGB, which produces white. A meson must have a primary color and its anticolor, also producing white.

![Figure 33.19](image.png)

Figure 33.19 The three quarks composing a baryon must be RGB, which add to white. The quark and antiquark composing a meson must be a color and anticolor, here \(R \bar{R}\) also adding to white. The force between systems that have color is so great that they can neither be separated nor exist as colored.

Why must hadrons be white? The color scheme is intentionally devised to explain why baryons have three quarks and mesons have a quark and an antiquark. Quark color is thought to be similar to charge, but with more values. An ion, by analogy, exerts much stronger forces than a neutral molecule. When the color of a combination of quarks is white, it is like a neutral atom. The forces a white particle exerts are like the polarization forces in molecules, but in hadrons these leftovers are the strong nuclear force. When a combination of quarks has color other than white, it exerts extremely large forces—even larger than the strong force—and perhaps cannot be stable or permanently separated. This is part of the theory of quark confinement, which explains how quarks can exist and yet never be isolated or directly observed. Finally, an extra quantum number with three values (like those we assign to color) is necessary for quarks to obey the Pauli exclusion principle. Particles such as the \(\Omega^-\), which is composed of three strange quarks, \(sss\), and the \(\Delta^{++}\), which is three up quarks, \(uuu\), can exist because the quarks have different colors and do not have the same quantum numbers. Color is consistent with all observations and is now widely accepted. Quark theory including color is called quantum chromodynamics (QCD), also named by Gell-Mann.

The Three Families

Fundamental particles are thought to be one of three types—leptons, quarks, or carrier particles. Each of those three types is further divided into three analogous families as illustrated in Figure 33.20. We have examined leptons and quarks in some detail. Each has six members (and their six antiparticles) divided into three analogous families. The first family is normal matter, of which most things are composed. The second is exotic, and the third more exotic and more massive than the second. The only stable particles are in the first family, which also has unstable members.

Always searching for symmetry and similarity, physicists have also divided the carrier particles into three families, omitting the graviton. Gravity is special among the four forces in that it affects the space and time in which the other forces exist and is proving most difficult to include in a Theory of Everything or TOE (to stub the pretension of such a theory). Gravity is thus often set apart. It is not certain that there is meaning in the groupings shown in Figure 33.20, but the analogies are tempting. In the past, we have been able to make significant advances by looking for analogies and patterns, and this is an example of one under current scrutiny. There are connections between the families of leptons, in that the \(\tau\) decays into the \(\mu\) and the \(\mu\) into the \(e\). Similarly for quarks, the higher families eventually decay into the lowest, leaving only \(u\) and \(d\) quarks. We have long sought connections between the forces in nature. Since these are carried by particles, we will explore connections between gluons, \(W^\pm\) and \(Z^0\), and photons as part of the search for unification of forces discussed in GUTs: The Unification of Forces..
**33.6 GUTs: The Unification of Forces**

Present quests to show that the four basic forces are different manifestations of a single unified force follow a long tradition. In the 19th century, the distinct electric and magnetic forces were shown to be intimately connected and are now collectively called the electromagnetic force. More recently, the weak nuclear force has been shown to be connected to the electromagnetic force in a manner suggesting that a theory may be constructed in which all four forces are unified. Certainly, there are similarities in how forces are transmitted by the exchange of carrier particles, and the carrier particles themselves (the gauge bosons in Table 33.2) are also similar in important ways. The analogy to the unification of electric and magnetic forces is quite good—the four forces are distinct under normal circumstances, but there are hints of connections even on the atomic scale, and there may be conditions under which the forces are intimately related and even indistinguishable. The search for a correct theory linking the forces, called the Grand Unified Theory (GUT), is explored in this section in the realm of particle physics. Frontiers of Physics expands the story in making a connection with cosmology, on the opposite end of the distance scale.

Figure 33.21 is a Feynman diagram showing how the weak nuclear force is transmitted by the carrier particle $Z^0$, similar to the diagrams in Figure 33.5 and Figure 33.6 for the electromagnetic and strong nuclear forces. In the 1960s, a gauge theory, called electroweak theory, was developed by Steven Weinberg, Sheldon Glashow, and Abdus Salam and proposed that the electromagnetic and weak forces are identical at sufficiently high energies. One of its predictions, in addition to describing both electromagnetic and weak force phenomena, was the existence of the $W^+$, $W^-$, and $Z^0$ carrier particles. Not only were three particles having spin 1 predicted, the mass of the $W^+$ and $W^-$ was predicted to be $81 \text{ GeV}/c^2$, and that of the $Z^0$ was predicted to be $90 \text{ GeV}/c^2$. (Their masses had to be about 1000 times that of the pion, or about $100 \text{ GeV}/c^2$, since the range of the weak force is about 1000 times less than the strong force carried by virtual pions.) In 1983, these carrier particles were observed at CERN with the predicted characteristics, including masses having the predicted values as seen in Table 33.2. This was another triumph of particle theory and experimental effort, resulting in the 1984 Nobel Prize to the experiment’s group leaders Carlo Rubbia and Simon van der Meer. Theorists Weinberg, Glashow, and Salam had already been honored with the 1979 Nobel Prize for other aspects of electroweak theory.

Although the weak nuclear force is very short ranged ($< 10^{-18} \text{ m}$, as indicated in Table 33.1), its effects on atomic levels can be measured given the extreme precision of modern techniques. Since electrons spend some time in the nucleus, their energies are affected, and spectra can even indicate new aspects of the weak force, such as the possibility of other carrier particles. So systems many orders of magnitude larger than the range of the weak force supply evidence of electroweak unification in addition to evidence found at the particle scale.
Gluons \((g)\) are the proposed carrier particles for the strong nuclear force, although they are not directly observed. Like quarks, gluons may be confined to systems having a total color of white. Less is known about gluons than the fact that they are the carriers of the weak and certainly of the electromagnetic force. QCD theory calls for eight gluons, all massless and all spin 1. Six of the gluons carry a color and an anticolor, while two do not carry color, as illustrated in Figure 33.22(a). There is indirect evidence of the existence of gluons in nucleons. When high-energy electrons are scattered from nucleons and evidence of quarks is seen, the momenta of the quarks are smaller than they would be if there were no gluons. That means that the gluons carrying force between quarks also carry some momentum, inferred by the already indirect quark momentum measurements. At any rate, the gluons carry color charge and can change the colors of quarks when exchanged, as seen in Figure 33.22(b). In the figure, a red down quark interacts with a green strange quark by sending it a gluon. That gluon carries red away from the down quark and leaves it green, because it is an \(RG\) (red-antigreen) gluon. (Taking antigreen away leaves you green.) Its antigreenness kills the green in the strange quark, and its redness turns the quark red.

The strong force is complicated, since observable particles that feel the strong force (hadrons) contain multiple quarks. Figure 33.23 shows the quark and gluon details of pion exchange between a proton and a neutron as illustrated earlier in Figure 33.3 and Figure 33.6. The quarks within the proton and neutron move along together exchanging gluons, until the proton and neutron get close together. As the \(u\) quark leaves the proton, a gluon creates a pair of virtual particles, a \(d\) quark and a \(\bar{d}\) antiquark. The \(d\) quark stays behind and the proton turns into a neutron, while the \(u\) and \(\bar{d}\) move together as a \(\pi^+\) \((\text{Table 33.4 confirms the } u\bar{d} \text{ composition for the } \pi^+).\) The \(\bar{d}\) annihilates a \(d\) quark in the neutron, the \(u\) joins the neutron, and the neutron becomes a proton. A pion is exchanged and a force is transmitted.

It is beyond the scope of this text to go into more detail on the types of quark and gluon interactions that underlie the observable particles, but the theory \((\text{quantum chromodynamics or QCD})\) is very self-consistent. So successful have QCD and the electroweak theory been that, taken together, they are called the \text{Standard Model}. Advances in knowledge are expected to modify, but not overthrow, the Standard Model of particle physics and forces.

**Making Connections: Unification of Forces**

Grand Unified Theory (GUT) is successful in describing the four forces as distinct under normal circumstances, but connected in fundamental ways. Experiments have verified that the weak and electromagnetic force become identical at very small distances and provide the GUT description of the carrier particles for the forces. GUT predicts that the other forces become identical under conditions so extreme that they cannot be tested in the laboratory, although there may be lingering evidence of them in the evolution of the universe. GUT is also successful in describing a system of carrier particles for all four forces, but there is much to be done, particularly in the realm of gravity.
How can forces be unified? They are definitely distinct under most circumstances, for example, being carried by different particles and having greatly different strengths. But experiments show that at extremely small distances, the strengths of the forces begin to become more similar. In fact, electroweak theory’s prediction of the $W^+$, $W^-$, and $Z^0$ carrier particles was based on the strengths of the two forces being identical at extremely small distances as seen in Figure 33.24. As discussed in case of the creation of virtual particles for extremely short times, the small distances or short ranges correspond to the large masses of the carrier particles and the correspondingly large energies needed to create them. Thus, the energy scale on the horizontal axis of Figure 33.24 corresponds to smaller and smaller distances, with 100 GeV corresponding to approximately, $10^{-18}$ m for example. At that distance, the strengths of the EM and weak forces are the same. To test physics at that distance, energies of about 100 GeV must be put into the system, and that is sufficient to create and release the $W^+$, $W^-$, and $Z^0$ carrier particles. At those and higher energies, the masses of the carrier particles becomes less and less relevant, and the $Z^0$ in particular resembles the massless, chargeless, spin 1 photon. In fact, there is enough energy when things are pushed to even smaller distances to transform the, and $Z^0$ into massless carrier particles more similar to photons and gluons. These have not been observed experimentally, but there is a prediction of an associated particle called the Higgs boson. The mass of this particle is not predicted with nearly the certainty with which the mass of the $W^+$, $W^-$, and $Z^0$ particles were predicted, but it was hoped that the Higgs boson could be observed at the now-canceled Superconducting Super Collider (SSC). Ongoing experiments at the Large Hadron Collider at CERN have presented some evidence for a Higgs boson with a mass of 125 GeV, and there is a possibility of a direct discovery during 2012. The existence of this more massive particle would give validity to the theory that the carrier particles are identical under certain circumstances.

The small distances and high energies at which the electroweak force becomes identical with the strong nuclear force are not reachable with any conceivable human-built accelerator. At energies of about $10^{14}$ GeV (16,000 J per particle), distances of about $10^{-30}$ m can be probed. Such energies are needed to test theory directly, but these are about $10^{10}$ higher than the proposed giant SSC would have had, and the distances are about $10^{-12}$ smaller than any structure we have direct knowledge of. This would be the realm of various GUTs, of which there are many since there is no constraining evidence at these energies and distances. Past experience has shown that any time you probe so many orders of magnitude further (here, about $10^{12}$ ), you find the unexpected. Even more extreme are the energies and distances at which gravity is thought to unify with the other forces in a TOE. Most speculative and least constrained by experiment are TOEs, one of which is called Superstring theory. Superstrings are entities that are $10^{-35}$ m in scale and act like one-dimensional oscillating strings and are also proposed to underlie all particles, forces, and space itself.

At the energy of GUTs, the carrier particles of the weak force would become massless and identical to gluons. If that happens, then both lepton and baryon conservation would be violated. We do not see such violations, because we do not encounter such energies. However, there is a tiny probability that, at ordinary energies, the virtual particles that violate the conservation of baryon number may exist for extremely small amounts of time (corresponding to very small ranges). All GUTs thus predict that the proton should be unstable, but would decay with an extremely long lifetime of about $10^{31}$ y. The predicted decay mode is

$$p \rightarrow \pi^0 + e^+,$$

(33.11)

which violates both conservation of baryon number and electron family number. Although $10^{31}$ y is an extremely long time (about $10^{21}$ times the age of the universe), there are a lot of protons, and detectors have been constructed to look for the proposed decay mode as seen in Figure 33.25. It is somewhat comforting that proton decay has not been detected, and its experimental lifetime is now greater than $5\times10^{32}$ y. This does not prove GUTs wrong, but it does place greater constraints on the theories, benefiting theorists in many ways.

From looking increasingly inward at smaller details for direct evidence of electroweak theory and GUTs, we turn around and look to the universe for evidence of the unification of forces. In the 1920s, the expansion of the universe was discovered. Thinking backward in time, the universe must once have been very small, dense, and extremely hot. At a tiny fraction of a second after the fabled Big Bang, forces would have been unified and may have left their fingerprint on the existing universe. This, one of the most exciting forefronts of physics, is the subject of Frontiers of Physics.
baryon number: a conserved physical quantity that is zero for mesons and leptons and $\pm 1$ for baryons and antibaryons, respectively

baryons: hadrons that always decay to another baryon

boson: particle with zero or an integer value of intrinsic spin

bottom: a quark flavor

charm: a quark flavor, which is the counterpart of the strange quark

colliding beams: head-on collisions between particles moving in opposite directions

color: a quark flavor

conservation of total baryon number: a general rule based on the observation that the total number of nucleons was always conserved in nuclear reactions and decays

conservation of total electron family number: a general rule stating that the total electron family number stays the same through an interaction

conservation of total muon family number: a general rule stating that the total muon family number stays the same through an interaction

cyclotron: accelerator that uses fixed-frequency alternating electric fields and fixed magnets to accelerate particles in a circular spiral path

down: the second-lightest of all quarks

electron family number: the number $\pm 1$ that is assigned to all members of the electron family, or the number 0 that is assigned to all particles not in the electron family

electroweak theory: theory showing connections between EM and weak forces

Feynman diagram: a graph of time versus position that describes the exchange of virtual particles between subatomic particles

fermion: particle with a half-integer value of intrinsic spin

flavors: quark type

fundamental particle: particle with no substructure

gauge boson: particle that carries one of the four forces

gluons: exchange particles, analogous to the exchange of photons that gives rise to the electromagnetic force between two charged particles

gluons: eight proposed particles which carry the strong force

grand unified theory: theory that shows unification of the strong and electroweak forces

Higgs boson: a massive particle that, if observed, would give validity to the theory that carrier particles are identical under certain circumstances
hadrons: particles that feel the strong nuclear force
leptons: particles that do not feel the strong nuclear force
linear accelerator: accelerator that accelerates particles in a straight line
meson: particle whose mass is intermediate between the electron and nucleon masses
meson: hadrons that can decay to leptons and leave no hadrons
muon family number: the number $\pm 1$ that is assigned to all members of the muon family, or the number 0 that is assigned to all particles not in the muon family
particle physics: the study of and the quest for those truly fundamental particles having no substructure
pion: particle exchanged between nucleons, transmitting the force between them
quantum chromodynamics: quark theory including color
quantum chromodynamics: the governing theory of connecting quantum number color to gluons
quantum electrodynamics: the theory of electromagnetism on the particle scale
quark: an elementary particle and a fundamental constituent of matter
standard model: combination of quantum chromodynamics and electroweak theory
strangeness: a physical quantity assigned to various particles based on decay systematics
strange: the third lightest of all quarks
superstring theory: a theory of everything based on vibrating strings some $10^{-35}$ m in length
synchrotron radiation: radiation caused by a magnetic field accelerating a charged particle perpendicular to its velocity
synchrotron: a version of a cyclotron in which the frequency of the alternating voltage and the magnetic field strength are increased as the beam particles are accelerated
tau family number: the number $\pm 1$ that is assigned to all members of the tau family, or the number 0 that is assigned to all particles not in the tau family
theory of quark confinement: explains how quarks can exist and yet never be isolated or directly observed
top: a quark flavor
up: the lightest of all quarks
Van de Graaff: early accelerator: simple, large-scale version of the electron gun
virtual particles: particles which cannot be directly observed but their effects can be directly observed

**Section Summary**

33.0 Introduction to Particle Physics
- Particle physics is the study of and the quest for those truly fundamental particles having no substructure.

33.1 The Yukawa Particle and the Heisenberg Uncertainty Principle Revisited
- Yukawa’s idea of virtual particle exchange as the carrier of forces is crucial, with virtual particles being formed in temporary violation of the conservation of mass-energy as allowed by the Heisenberg uncertainty principle.

33.2 The Four Basic Forces
- The four basic forces and their carrier particles are summarized in the Table 33.1.
- Feynman diagrams are graphs of time versus position and are highly useful pictorial representations of particle processes.
- The theory of electromagnetism on the particle scale is called quantum electrodynamics (QED).

33.3 Accelerators Create Matter from Energy
- A variety of particle accelerators have been used to explore the nature of subatomic particles and to test predictions of particle theories.
- Modern accelerators used in particle physics are either large synchrotrons or linear accelerators.
- The use of colliding beams makes much greater energy available for the creation of particles, and collisions between matter and antimatter allow a greater range of final products.

33.4 Particles, Patterns, and Conservation Laws
33.5 Quarks: Is That All There Is?
- Hadrons are thought to be composed of quarks, with baryons having three quarks and mesons having a quark and an antiquark.
- The characteristics of the six quarks and their antiquark counterparts are given in Table 33.3, and the quark compositions of certain hadrons are given in Table 33.4.
- Indirect evidence for quarks is very strong, explaining all known hadrons and their quantum numbers, such as strangeness, charm, topness, and bottomness.
- Quarks come in six flavors and three colors and occur only in combinations that produce white.
- Fundamental particles have no further substructure, not even a size beyond their de Broglie wavelength.
- There are three types of fundamental particles—leptons, quarks, and carrier particles. Each type is divided into three analogous families as indicated in Figure 33.20.

33.6 GUTs: The Unification of Forces
- Attempts to show unification of the four forces are called Grand Unified Theories (GUTs) and have been partially successful, with connections proven between EM and weak forces in electroweak theory.
- The strong force is carried by eight proposed particles called gluons, which are intimately connected to a quantum number called color—their governing theory is thus called quantum chromodynamics (QCD). Taken together, QCD and the electroweak theory are widely accepted as the Standard Model of particle physics.
- Unification of the strong force is expected at such high energies that it cannot be directly tested, but it may have observable consequences in the as-yet unobserved decay of the proton and topics to be discussed in the next chapter. Although unification of forces is generally anticipated, much remains to be done to prove its validity.

Conceptual Questions

33.3 Accelerators Create Matter from Energy
1. The total energy in the beam of an accelerator is far greater than the energy of the individual beam particles. Why isn’t this total energy available to create a single extremely massive particle?
2. Synchrotron radiation takes energy from an accelerator beam and is related to acceleration. Why would you expect the problem to be more severe for electron accelerators than proton accelerators?
3. What two major limitations prevent us from building high-energy accelerators that are physically small?
4. What are the advantages of colliding-beam accelerators? What are the disadvantages?

33.4 Particles, Patterns, and Conservation Laws
5. Large quantities of antimatter isolated from normal matter should behave exactly like normal matter. An antiatom, for example, composed of positrons, antiprotons, and antineutrons should have the same atomic spectrum as its matter counterpart. Would you be able to tell it is antimatter by its emission of antiphotons? Explain briefly.
6. Massless particles are not only neutral, they are chargeless (unlike the neutron). Why is this so?
7. Massless particles must travel at the speed of light, while others cannot reach this speed. Why are all massless particles stable? If evidence is found that neutrinos spontaneously decay into other particles, would this imply they have mass?
8. When a star erupts in a supernova explosion, huge numbers of electron neutrinos are formed in nuclear reactions. Such neutrinos from the 1987A supernova in the relatively nearby Magellanic Cloud were observed within hours of the initial brightening, indicating they traveled to earth at approximately the speed of light. Explain how this data can be used to set an upper limit on the mass of the neutrino, noting that if the mass is small the neutrinos could travel very close to the speed of light and have a reasonable energy (on the order of MeV).
9. Theorists have had spectacular success in predicting previously unknown particles. Considering past theoretical triumphs, why should we bother to perform experiments?
10. What lifetime do you expect for an antineutron isolated from normal matter?
11. Why does the \( \eta^0 \) meson have such a short lifetime compared to most other mesons?
12. (a) Is a hadron always a baryon?
   (b) Is a baryon always a hadron?
   (c) Can an unstable baryon decay into a meson, leaving no other baryon?
13. Explain how conservation of baryon number is responsible for conservation of total atomic mass (total number of nucleons) in nuclear decay and reactions.

33.5 Quarks: Is That All There Is?
14. The quark flavor change \( d \rightarrow u \) takes place in \( \beta^- \) decay. Does this mean that the reverse quark flavor change \( u \rightarrow d \) takes place in \( \beta^+ \) decay? Justify your response by writing the decay in terms of the quark constituents, noting that it looks as if a proton is converted into a neutron in \( \beta^+ \) decay.

15. Explain how the weak force can change strangeness by changing quark flavor.

16. Beta decay is caused by the weak force, as are all reactions in which strangeness changes. Does this imply that the weak force can change quark flavor? Explain.

17. Why is it easier to see the properties of the \( c, b, \) and \( t \) quarks in mesons having composition \( W^- \) or \( t \tilde{t} \) rather than in baryons having a mixture of quarks, such as \( udb \)?

18. How can quarks, which are fermions, combine to form bosons? Why must an even number combine to form a boson? Give one example by stating the quark substructure of a boson.

19. What evidence is cited to support the contention that the gluon force between quarks is greater than the strong nuclear force between hadrons? How is this related to color? Is it also related to quark confinement?

20. Discuss how we know that \( \pi \)-mesons \( (\pi^+, \pi, \pi^0) \) are not fundamental particles and are not the basic carriers of the strong force.

21. An antibaryon has three antiquarks with colors \( \tilde{R}\tilde{G}\tilde{B} \). What is its color?

22. Suppose leptons are created in a reaction. Does this imply the weak force is acting? (for example, consider \( \beta^- \) decay.)

23. How can the lifetime of a particle indicate that its decay is caused by the strong nuclear force? How can a change in strangeness imply which force is responsible for a reaction? What does a change in quark flavor imply about the force that is responsible?

24. (a) Do all particles having strangeness also have at least one strange quark in them? (b) Do all hadrons with a strange quark also have nonzero strangeness?

25. The sigma-zero particle decays mostly via the reaction \( \Sigma^0 \rightarrow \Lambda^0 + \gamma \). Explain how this decay and the respective quark compositions imply that the \( \Sigma^0 \) is an excited state of the \( \Lambda^0 \).

26. What do the quark compositions and other quantum numbers imply about the relationships between the \( \Delta^+ \) and the proton? The \( \Delta^0 \) and the neutron?

27. Discuss the similarities and differences between the photon and the \( Z^0 \) in terms of particle properties, including forces felt.

28. Identify evidence for electroweak unification.

29. The quarks in a particle are confined, meaning individual quarks cannot be directly observed. Are gluons confined as well? Explain

33.6 GUTs: The Unification of Forces

30. If a GUT is proven, and the four forces are unified, it will still be correct to say that the orbit of the moon is determined by the gravitational force. Explain why.

31. If the Higgs boson is discovered and found to have mass, will it be considered the ultimate carrier of the weak force? Explain your response.

32. Gluons and the photon are massless. Does this imply that the \( W^+, W^- \), and \( Z^0 \) are the ultimate carriers of the weak force?
33.1 The Yukawa Particle and the Heisenberg Uncertainty Principle Revisited

1. A virtual particle having an approximate mass of $10^{14}$ GeV/c$^2$ may be associated with the unification of the strong and electroweak forces. For what length of time could this virtual particle exist (in temporary violation of the conservation of mass-energy as allowed by the Heisenberg uncertainty principle)?

2. Calculate the mass in GeV/c$^2$ of a virtual carrier particle that has a range limited to $10^{-30}$ m by the Heisenberg uncertainty principle. Such a particle might be involved in the unification of the strong and electroweak forces.

3. Another component of the strong nuclear force is transmitted by the exchange of virtual $K$-mesons. Taking $K$-mesons to have an average mass of 495 MeV/c$^2$, what is the approximate range of this component of the strong force?

33.2 The Four Basic Forces

4. (a) Find the ratio of the strengths of the weak and electromagnetic forces under ordinary circumstances.

(b) What does that ratio become under circumstances in which the forces are unified?

5. The ratio of the strong to the weak force and the ratio of the strong force to the electromagnetic force become 1 under circumstances where they are unified. What are the ratios of the strong force to those two forces under normal circumstances?

33.3 Accelerators Create Matter from Energy

6. At full energy, protons in the 2.00-km-diameter Fermilab synchrotron travel at nearly the speed of light, since their energy is about 1000 times their rest mass energy.

(a) How long does it take for a proton to complete one trip around?

(b) How many times per second will it pass through the target area?

7. Suppose a $W^-$ created in a bubble chamber lives for $5.00 \times 10^{-25}$ s. What distance does it move in this time if it is traveling at 0.990 c? Since this distance is too short to make a track, the presence of the $W^-$ must be inferred from its decay products. Note that the time is longer than the given $W^-$ lifetime, which can be due to the statistical nature of decay or time dilation.

8. What length track does a $\pi^+$ traveling at 0.100 c leave in a bubble chamber if it is created there and lives for $2.60 \times 10^{-8}$ s? (Those moving faster or living longer may escape the detector before decaying.)

9. The 3.20-km-long SLAC produces a beam of 50.0-GeV electrons. If there are 15,000 accelerating tubes, what average voltage must be across the gaps between them to achieve this energy?

10. Because of energy loss due to synchrotron radiation in the LHC at CERN, only 5.00 MeV is added to the energy of each proton during each revolution around the main ring. How many revolutions are needed to produce 7.00-TeV (7000 GeV) protons, if they are injected with an initial energy of 8.00 GeV?

11. A proton and an antiproton collide head-on, each having a kinetic energy of 7.00 TeV (such as in the LHC at CERN). How much collision energy is available, taking into account the annihilation of the two masses? (Note that this is not significantly greater than the extremely relativistic kinetic energy.)

12. When an electron and positron collide at the SLAC facility, they each have 50.0 GeV kinetic energies. What is the total collision energy available, taking into account the annihilation energy? Note that the annihilation energy is insignificant, because the electrons are highly relativistic.

33.4 Particles, Patterns, and Conservation Laws

13. The $\pi^0$ is its own antiparticle and decays in the following manner: $\pi^0 \rightarrow \gamma + \gamma$. What is the energy of each $\gamma$ ray if the $\pi^0$ is at rest when it decays?

14. The primary decay mode for the negative pion is $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$. What is the energy release in MeV in this decay?

15. The mass of a theoretical particle that may be associated with the unification of the electroweak and strong forces is $10^{14}$ GeV/c$^2$. (a) How many proton masses is this? (b) How many electron masses is this? (This indicates how extremely relativistic the accelerator would have to be in order to make the particle, and how large the relativistic quantity $\gamma$ would have to be.)

16. The decay mode of the negative muon is $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$. (a) Find the energy released in MeV. (b) Verify that charge and lepton family numbers are conserved.

17. The decay mode of the positive tau is $\tau^+ \rightarrow \mu^+ + \nu_\mu + \bar{\nu}_\tau$. (a) What energy is released? (b) Verify that charge and lepton family numbers are conserved. (c) The $\tau^+$ is the antiparticle of the $\tau^-$. Verify that all the decay products of the $\tau^+$ are the antiparticles of those in the decay of the $\tau^-$ given in the text.

18. The principal decay mode of the sigma zero is $\Sigma^0 \rightarrow \Lambda^0 + \gamma$. (a) What energy is released? (b) Considering the quark structure of the two baryons, does it appear that the $\Sigma^0$ is an excited state of the $\Lambda^0$? (c) Verify that strangeness, charge, and baryon number are conserved in the decay. (d) Considering the preceding and the short lifetime, can the weak force be responsible? State why or why not.

19. (a) What is the uncertainty in the energy released in the decay of a $\pi^0$ due to its short lifetime? (b) What fraction of the decay energy is this, noting that the decay mode is $\pi^0 \rightarrow \gamma + \gamma$ (so that all the $\pi^0$ mass is destroyed)? (20. (a) What is the uncertainty in the energy released in the decay of a $\tau^-$ due to its short lifetime? (b) Is the uncertainty in this energy greater than or less than the uncertainty in the mass of the tau neutrino? Discuss the source of the uncertainty.

33.5 Quarks: Is That All There Is?

21. (a) Verify from its quark composition that the $\Delta^+$ particle could be an excited state of the proton. (b) There is a spread of about 100 MeV in the decay energy of the $\Delta^+$, interpreted as uncertainty due to its short lifetime. What is its approximate lifetime? (c) Does its decay proceed via the strong or weak force?

22. Accelerators such as the Triangle Universities Meson Facility (TRIUMF) in British Columbia produce secondary beams of pions by...
having an intense primary proton beam strike a target. Such “meson factories” have been used for many years to study the interaction of pions with nuclei and, hence, the strong nuclear force. One reaction that occurs is \( \pi^+ + p \to \Delta^{++} \to \pi^+ + p \), where the \( \Delta^{++} \) is a very short-lived particle. The graph in Figure 33.26 shows the probability of this reaction as a function of energy. The width of the bump is the uncertainty in energy due to the short lifetime of the \( \Delta^{++} \).

(a) Find this lifetime.
(b) Verify from the quark composition of the particles that this reaction annihilates and then re-creates a \( d \) quark and a \( \bar{d} \) antiquark by writing the reaction and decay in terms of quarks.
(c) Draw a Feynman diagram of the production and decay of the \( \Delta^{++} \) showing the individual quarks involved.

23. The reaction \( \pi^+ + p \to \Delta^{++} \) (described in the preceding problem) takes place via the strong force. (a) What is the baryon number of the \( \Delta^{++} \) particle?
(b) Draw a Feynman diagram of the reaction showing the individual quarks involved.

24. One of the decay modes of the omega minus is \( \Omega^- \to \Sigma^0 + \pi^- \).
(a) What is the change in strangeness?
(b) Verify that baryon number and charge are conserved, while lepton numbers are unaffected.
(c) Write the equation in terms of the constituent quarks, indicating that the weak force is responsible.

25. Repeat the previous problem for the decay mode \( \Omega^- \to \Lambda^0 + K^- \).

26. One decay mode for the eta-zero meson is \( \eta^0 \to \gamma + \gamma \).
(a) Find the energy released.
(b) What is the uncertainty in the energy due to the short lifetime?
(c) Write the decay in terms of the constituent quarks.
(d) Verify that baryon number, lepton numbers, and charge are conserved.

27. One decay mode for the eta-zero meson is \( \eta^0 \to \pi^0 + \pi^0 \).
(a) Write the decay in terms of the quark constituents.
(b) How much energy is released?
(c) What is the ultimate release of energy, given the decay mode for the pi zero is \( \pi^0 \to \gamma + \gamma \)?
create a shower of radiation from all the particles they create by striking a nucleus in the atmosphere as seen in the figure given below. Suppose a cosmic ray particle having an energy of \(10^{10} \text{ GeV}\) converts its energy into particles with masses averaging \(200 \text{ MeV}/c^2\). (a) How many particles are created? (b) If the particles rain down on a \(1.00-\text{km}^2\) area, how many particles are there per square meter?

Figure 33.27 An extremely energetic cosmic ray creates a shower of particles on earth. The energy of these rare cosmic rays can approach a joule (about \(10^{10} \text{ GeV}\) ) and, after multiple collisions, huge numbers of particles are created from this energy. Cosmic ray showers have been observed to extend over many square kilometers.

44. Integrated Concepts
Assuming conservation of momentum, what is the energy of each \(\gamma\) ray produced in the decay of a neutral at rest pion, in the reaction \(\pi^0 \rightarrow \gamma + \gamma\)?

45. Integrated Concepts
What is the wavelength of a 50-GeV electron, which is produced at SLAC? This provides an idea of the limit to the detail it can probe.

46. Integrated Concepts
(a) Calculate the relativistic quantity \(\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}\) for 1.00-TeV protons produced at Fermilab. (b) If such a proton created a \(\pi^+\) having the same speed, how long would its life be in the laboratory? (c) How far could it travel in this time?

47. Integrated Concepts
The primary decay mode for the negative pion is \(\pi^- \rightarrow \mu^- + \bar{\nu}_\mu\). (a) What is the energy release in MeV in this decay? (b) Using conservation of momentum, how much energy does each of the decay products receive, given the \(\pi^-\) is at rest when it decays? You may assume the muon antineutrino is massless and has momentum \(p = E/c\), just like a photon.

48. Integrated Concepts
Plans for an accelerator that produces a secondary beam of \(K\)-mesons to scatter from nuclei, for the purpose of studying the strong force, call for them to have a kinetic energy of 500 MeV. (a) What would the relativistic quantity \(\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}\) be for these particles? (b) How long would their average lifetime be in the laboratory? (c) How far could they travel in this time?

49. Integrated Concepts
Suppose you are designing a proton decay experiment and you can detect 50 percent of the proton decays in a tank of water. (a) How many kilograms of water would you need to see one decay per month, assuming a lifetime of \(10^{31}\) years? (b) How many cubic meters of water is this? (c) If the actual lifetime is \(10^{33}\) years, how long would you have to wait on an average to see a single proton decay?

50. Construct Your Own Problem
Consider an ultrahigh-energy cosmic ray entering the Earth’s atmosphere (some have energies approaching a joule). Construct a problem in which you calculate the energy of the particle based on the number of particles in an observed cosmic ray shower. Among the things to consider are the average mass of the shower particles, the average number per square meter, and the extent (number of square meters covered) of the shower. Express the energy in eV and joules.

51. Construct Your Own Problem
Consider an ultrahigh-energy cosmic ray that enters the Earth’s atmosphere. Construct a problem in which you calculate the energy of the particle based on the number of particles in an observed cosmic ray shower. Among the things to consider are the average mass of the shower particles, the average number per square meter, and the extent of the shower. Express the energy in eV and joules.

52. Construct Your Own Problem
Consider a detector needed to observe the proposed, but extremely rare, decay of an electron. Construct a problem in which you calculate the amount of matter needed in the detector to be able to observe the decay, assuming that it has a signature that is clearly identifiable. Among the things to consider are the estimated half life (long for rare events), and the number of decays per unit time that you wish to observe, as well as the number of electrons in the detector substance.
Learning Objectives

34.1. Cosmology and Particle Physics
   • Discuss the expansion of the universe.
   • Explain the Big Bang.

34.2. General Relativity and Quantum Gravity
   • Explain the effect of gravity on light.
   • Discuss black hole.
   • Explain quantum gravity.

34.3. Superstrings
   • Define Superstring theory.
   • Explain the relationship between Superstring theory and the Big Bang.

34.4. Dark Matter and Closure
   • Discuss the existence of dark matter.
   • Explain neutrino oscillations and their consequences.

34.5. Complexity and Chaos
   • Explain complex systems.
   • Discuss chaotic behavior of different systems.

34.6. High-temperature Superconductors
   • Identify superconductors and their uses.
   • Discuss the need for a high-\(T_c\) superconductor.

34.7. Some Questions We Know to Ask
   • Identify sample questions to be asked on the largest scales.
   • Identify sample questions to be asked on the intermediate scale.
   • Identify sample questions to be asked on the smallest scales.

Introduction to Frontiers of Physics

Frontiers are exciting. There is mystery, surprise, adventure, and discovery. The satisfaction of finding the answer to a question is made keener by the fact that the answer always leads to a new question. The picture of nature becomes more complete, yet nature retains its sense of mystery and never loses its ability to awe us. The view of physics is beautiful looking both backward and forward in time. What marvelous patterns we have discovered. How clever nature seems in its rules and connections. How awesome. And we continue looking ever deeper and ever further, probing the basic structure of matter, energy, space, and time and wondering about the scope of the universe, its beginnings and future.

You are now in a wonderful position to explore the frontiers of physics, both the new discoveries and the unanswered questions. With the concepts, qualitative and quantitative, the problem-solving skills, the feeling for connections among topics, and all the rest you have mastered, you can more deeply appreciate and enjoy the brief treatments that follow. Years from now you will still enjoy the quest with an insight all the greater for your efforts.
34.1 Cosmology and Particle Physics

Look at the sky on some clear night when you are away from city lights. There you will see thousands of individual stars and a faint glowing background of millions more. The Milky Way, as it has been called since ancient times, is an arm of our galaxy of stars—the word galaxy coming from the Greek word galaxias, meaning milky. We know a great deal about our Milky Way galaxy and of the billions of other galaxies beyond its fringes. But they still provoke wonder and awe (see Figure 34.2). And there are still many questions to be answered. Most remarkable when we view the universe on the large scale is that once again explanations of its character and evolution are tied to the very small scale. Particle physics and the questions being asked about the very small scales may also have their answers in the very large scales.

Figure 34.2 Take a moment to contemplate these clusters of galaxies, photographed by the Hubble Space Telescope. Trillions of stars linked by gravity in fantastic forms, glowing with light and showing evidence of undiscovered matter. What are they like, these myriad stars? How did they evolve? What can they tell us of matter, energy, space, and time? (credit: NASA, ESA, K. Sharon (Tel Aviv University) and E. Ofek (Caltech))

As has been noted in numerous Things Great and Small vignettes, this is not the first time the large has been explained by the small and vice versa. Newton realized that the nature of gravity on Earth that pulls an apple to the ground could explain the motion of the moon and planets so much farther away. Minute atoms and molecules explain the chemistry of substances on a much larger scale. Decays of tiny nuclei explain the hot interior of the Earth. Fusion of nuclei likewise explains the energy of stars. Today, the patterns in particle physics seem to be explaining the evolution and character of the universe. And the nature of the universe has implications for unexplored regions of particle physics.

Cosmology is the study of the character and evolution of the universe. What are the major characteristics of the universe as we know them today? First, there are approximately $10^{11}$ galaxies in the observable part of the universe. An average galaxy contains more than $10^{11}$ stars, with our Milky Way galaxy being larger than average, both in its number of stars and its dimensions. Ours is a spiral-shaped galaxy with a diameter of about 100,000 light years and a thickness of about 2000 light years in the arms with a central bulge about 10,000 light years across. The Sun lies about 30,000 light years from the center near the galactic plane. There are significant clouds of gas, and there is a halo of less-dense regions of stars surrounding the main body. (See Figure 34.3.) Evidence strongly suggests the existence of a large amount of additional matter in galaxies that does not produce light—the mysterious dark matter we shall later discuss.
The Milky Way galaxy is typical of large spiral galaxies in its size, its shape, and the presence of gas and dust. We are fortunate to be in a location where we can see out of the galaxy and observe the vastly larger and fascinating universe around us. (a) Side view. (b) View from above. (c) The Milky Way as seen from Earth. (credits: (a) NASA, (b) Nick Risinger, (c) Andy)

Distances are great even within our galaxy and are measured in light years (the distance traveled by light in one year). The average distance between galaxies is on the order of a million light years, but it varies greatly with galaxies forming clusters such as shown in Figure 34.2. The Magellanic Clouds, for example, are small galaxies close to our own, some 160,000 light years from Earth. The Andromeda galaxy is a large spiral galaxy like ours and lies 2 million light years away. It is just visible to the naked eye as an extended glow in the Andromeda constellation. Andromeda is the closest large galaxy in our local group, and we can see some individual stars in it with our larger telescopes. The most distant known galaxy is 14 billion light years from Earth—a truly incredible distance. (See Figure 34.4.)
Consider the fact that the light we receive from these vast distances has been on its way to us for a long time. In fact, the time in years is the same as the distance in light years. For example, the Andromeda galaxy is 2 million light years away, so that the light now reaching us left it 2 million years ago. If we could be there now, Andromeda would be different. Similarly, light from the most distant galaxy left it 14 billion years ago. We have an incredible view of the past when looking great distances. We can try to see if the universe was different then—if distant galaxies are more tightly packed or have younger-looking stars, for example, than closer galaxies, in which case there has been an evolution in time. But the problem is that the uncertainties in our data are great. Cosmology is almost typified by these large uncertainties, so that we must be especially cautious in drawing conclusions. One consequence is that there are more questions than answers, and so there are many competing theories. Another consequence is that any hard data produce a major result. Discoveries of some importance are being made on a regular basis, the hallmark of a field in its golden age.

Perhaps the most important characteristic of the universe is that all galaxies except those in our local cluster seem to be moving away from us at speeds proportional to their distance from our galaxy. It looks as if a gigantic explosion, universally called the Big Bang, threw matter out some billions of years ago. This amazing conclusion is based on the pioneering work of Edwin Hubble (1889–1953), the American astronomer. In the 1920s, Hubble first demonstrated conclusively that other galaxies, many previously called nebulae or clouds of stars, were outside our own. He then found that all but the closest galaxies have a red shift in their hydrogen spectra that is proportional to their distance. The explanation is that there is a cosmological red shift due to the expansion of space itself. The photon wavelength is stretched in transit from the source to the observer. Double the distance, and the red shift is doubled. While this cosmological red shift is often called a Doppler shift, it is not—space itself is expanding. There is no center of expansion in the universe. All observers see themselves as stationary; the other objects in space appear to be moving away from them. Hubble was directly responsible for discovering that the universe was much larger than had previously been imagined and that it had this amazing characteristic of rapid expansion.

Universal expansion on the scale of galactic clusters (that is, galaxies at smaller distances are not uniformly receding from one another) is an integral part of modern cosmology. For galaxies farther away than about 50 Mly (50 million light years), the expansion is uniform with variations due to local motions of galaxies within clusters. A representative recession velocity \( v \) can be obtained from the simple formula

\[
    v = H_0d,
\]

where \( d \) is the distance to the galaxy and \( H_0 \) is the Hubble constant. The Hubble constant is a central concept in cosmology. Its value is determined by taking the slope of a graph of velocity versus distance, obtained from red shift measurements, such as shown in Figure 34.5. We shall use an approximate value of \( H_0 = 20 \text{ km/s} \cdot \text{Mly} \). Thus, \( v = H_0d \) is an average behavior for all but the closest galaxies. For example, a galaxy 100 Mly away (as determined by its size and brightness) typically moves away from us at a speed of \( v = (20 \text{ km/s} \cdot \text{Mly})(100 \text{ Mly}) = 2000 \text{ km/s} \). There can be variations in this speed due to so-called local motions or interactions with neighboring galaxies. Conversely, if a galaxy is found to be moving away from us at speed of 100,000 km/s based on its red shift, it is at a distance \( d = v / H_0 = (10,000 \text{ km/s}) / (20 \text{ km/s} \cdot \text{Mly}) = 5000 \text{ Mly} = 5 \text{ Gly} \text{ or } 5 \times 10^9 \text{ ly} \). This last calculation is approximate, because it assumes the expansion rate was the same 5 billion years ago as now. A similar calculation in Hubble’s measurement changed the notion that the universe is in a steady state.
Figure 34.5 This graph of red shift versus distance for galaxies shows a linear relationship, with larger red shifts at greater distances, implying an expanding universe. The slope gives an approximate value for the expansion rate. (credit: John Cub).

One of the most intriguing developments recently has been the discovery that the expansion of the universe may be faster now than in the past, rather than slowing due to gravity as expected. Various groups have been looking, in particular, at supernovas in moderately distant galaxies (less than 1 Gly) to get improved distance measurements. Those distances are larger than expected for the observed galactic red shifts, implying the expansion was slower when that light was emitted. This has cosmological consequences that are discussed in Dark Matter and Closure. The first results, published in 1999, are only the beginning of emerging data, with astronomy now entering a data-rich era.

Figure 34.6 shows how the recession of galaxies looks like the remnants of a gigantic explosion, the famous Big Bang. Extrapolating backward in time, the Big Bang would have occurred between 13 and 15 billion years ago when all matter would have been at a point. Questions instantly arise. What caused the explosion? What happened before the Big Bang? Was there a before, or did time start then? Will the universe expand forever, or will gravity reverse it into a Big Crunch? And is there other evidence of the Big Bang besides the well-documented red shifts?

The Russian-born American physicist George Gamow (1904–1968) was among the first to note that, if there was a Big Bang, the remnants of the primordial fireball should still be evident and should be blackbody radiation. Since the radiation from this fireball has been traveling to us since shortly after the Big Bang, its wavelengths should be greatly stretched. It will look as if the fireball has cooled in the billions of years since the Big Bang. Gamow and collaborators predicted in the late 1940s that there should be blackbody radiation from the explosion filling space with a characteristic temperature of about 7 K. Such blackbody radiation would have its peak intensity in the microwave part of the spectrum. (See Figure 34.7.) In 1964, Arno Penzias and Robert Wilson, two American scientists working with Bell Telephone Laboratories on a low-noise radio antenna, detected the radiation and eventually recognized it for what it is.

Figure 34.7(b) shows the spectrum of this microwave radiation that permeates space and is of cosmic origin. It is the most perfect blackbody spectrum known, and the temperature of the fireball remnant is determined from it to be $2.725 \pm 0.002K$. The detection of what is now called the cosmic microwave background (CMBR) was so important (generally considered as important as Hubble’s detection that the galactic red shift is proportional to distance) that virtually every scientist has accepted the expansion of the universe as fact. Penzias and Wilson shared the 1978 Nobel Prize in Physics for their discovery.
Figure 34.7 (a) The Big Bang is used to explain the present observed expansion of the universe. It was an incredibly energetic explosion some 10 to 20 billion years ago. After expanding and cooling, galaxies form inside the now-cold remnants of the primordial fireball. (b) The spectrum of cosmic microwave radiation is the most perfect blackbody spectrum ever detected. It is characteristic of a temperature of 2.725 K, the expansion-cooled temperature of the Big Bang’s remnant. This radiation can be measured coming from any direction in space not obscured by some other source. It is compelling evidence of the creation of the universe in a gigantic explosion, already indicated by galactic red shifts.

Making Connections: Cosmology and Particle Physics

There are many connections of cosmology—by definition involving physics on the largest scale—with particle physics—by definition physics on the smallest scale. Among these are the dominance of matter over antimatter, the nearly perfect uniformity of the cosmic microwave background, and the mere existence of galaxies.

Matter versus antimatter We know from direct observation that antimatter is rare. The Earth and the solar system are nearly pure matter. Space probes and cosmic rays give direct evidence—the landing of the Viking probes on Mars would have been spectacular explosions of mutual annihilation energy if Mars were antimatter. We also know that most of the universe is dominated by matter. This is proven by the lack of annihilation radiation coming to us from space, particularly the relative absence of 0.511-MeV γ rays created by the mutual annihilation of electrons and positrons. It seemed possible that there could be entire solar systems or galaxies made of antimatter in perfect symmetry with our matter-dominated systems. But the interactions between stars and galaxies would sometimes bring matter and antimatter together in large amounts. The annihilation radiation they would produce is simply not observed. Antimatter in nature is created in particle collisions and in $\beta^+$ decays, but only in small amounts that quickly annihilate, leaving almost pure matter surviving.

Particle physics seems symmetric in matter and antimatter. Why isn’t the cosmos? The answer is that particle physics is not quite perfectly symmetric in this regard. The decay of one of the neutral $K$-mesons, for example, preferentially creates more matter than antimatter. This is caused by a fundamental small asymmetry in the basic forces. This small asymmetry produced slightly more matter than antimatter in the early universe. If there was only one part in $10^9$ more matter (a small asymmetry), the rest would annihilate pair for pair, leaving nearly pure matter to form the stars and galaxies we see today. So the vast number of stars we observe may be only a tiny remnant of the original matter created in the Big Bang. Here at last we see a very real and important asymmetry in nature. Rather than be disturbed by an asymmetry, most physicists are impressed by how small it is. Furthermore, if the universe were completely symmetric, the mutual annihilation would be more complete, leaving far less matter to form us and the universe we know.
How can something so old have so few wrinkles? A troubling aspect of cosmic microwave background radiation (CMBR) was soon recognized. True, the CMBR verified the Big Bang, had the correct temperature, and had a blackbody spectrum as expected. But the CMBR was too smooth—it looked identical in every direction. Galaxies and other similar entities could not be formed without the existence of fluctuations in the primordial stages of the universe and so there should be hot and cool spots in the CMBR, nicknamed wrinkles, corresponding to dense and sparse regions of gas caused by turbulence or early fluctuations. Over time, dense regions would contract under gravity and form stars and galaxies. Why aren’t the fluctuations there? (This is a good example of an answer producing more questions.) Furthermore, galaxies are observed very far from us, so that they formed very long ago. The problem was to explain how galaxies could form so early and so quickly after the Big Bang if its remnant fingerprint is perfectly smooth. The answer is that if you look very closely, the CMBR is not perfectly smooth, only extremely smooth.

A satellite called the Cosmic Background Explorer (COBE) carried an instrument that made very sensitive and accurate measurements of the CMBR. In April of 1992, there was extraordinary publicity of COBE’s first results—there were small fluctuations in the CMBR. Further measurements were carried out by experiments including NASA’s Wilkinson Microwave Anisotropy Probe (WMAP), which launched in 2001. Data from WMAP provided a much more detailed picture of the CMBR fluctuations. (See Figure 34.7.) These amount to temperature fluctuations of only $200\mu K$ out of 2.7 K, better than one part in 1000. The WMAP experiment will be followed up by the European Space Agency’s Planck Surveyor, which launched in 2009.

Let us now examine the various stages of the overall evolution of the universe from the Big Bang to the present, illustrated in Figure 34.9. Note that scientific notation is used to encompass the many orders of magnitude in time, energy, temperature, and size of the universe. Going back in time, the two lines approach but do not cross (there is no zero on an exponential scale). Rather, they extend indefinitely in ever-smaller time intervals to some infinitesimal point.

Going back in time is equivalent to what would happen if expansion stopped and gravity pulled all the galaxies together, compressing and heating all matter. At a time long ago, the temperature and density were too high for stars and galaxies to exist. Before then, there was a time when the temperature was too great for atoms to exist. And farther back yet, there was a time when the temperature and density were so great that nuclei could not exist. Even farther back in time, the temperature was so high that average kinetic energy was great enough to create short-lived particles, and the density was high enough to make this likely. When we extrapolate back to the point of $W^\pm$ and $Z^{0}$ production (thermal energies reaching 1 TeV, or a temperature of about $10^{15}$ K), we reach the limits of what we know directly about particle physics. This is at a time about $10^{-12}$ s after
the Big Bang. While $10^{-12}$ s may seem to be negligibly close to the instant of creation, it is not. There are important stages before this time that are tied to the unification of forces. At those stages, the universe was at extremely high energies and average particle separations were smaller than we can achieve with accelerators. What happened in the early stages before $10^{-12}$ s is crucial to all later stages and is possibly discerned by observing present conditions in the universe. One of these is the smoothness of the CMBR.

Names are given to early stages representing key conditions. The stage before $10^{-11}$ s back to $10^{-34}$ s is called the **electroweak epoch**, because the electromagnetic and weak forces become identical for energies above about 100 GeV. As discussed earlier, theorists expect that the strong force becomes identical to and thus unified with the electroweak force at energies of about $10^{14}$ GeV. The average particle energy would be this great at $10^{-34}$ s after the Big Bang, if there are no surprises in the unknown physics at energies above about 1 TeV. At the immense energy of $10^{14}$ GeV (corresponding to a temperature of about $10^{26}$ K), the $W^\pm$ and $Z^0$ carrier particles would be transformed into massless gauge bosons to accomplish the unification. Before $10^{-34}$ s back to about $10^{-43}$ s, we have Grand Unification in the **GUT epoch**, in which all forces except gravity are identical. At $10^{-43}$ s, the average energy reaches the immense $10^{19}$ GeV needed to unify gravity with the other forces in TOE, the Theory of Everything. Before that time is the **TOE epoch**, but we have almost no idea as to the nature of the universe then, since we have no workable theory of quantum gravity. We call the hypothetical unified force **superforce**.

Now let us imagine starting at TOE and moving forward in time to see what type of universe is created from various events along the way. As temperatures and average energies decrease with expansion, the universe reaches the stage where average particle separations are large enough to see differences between the strong and electroweak forces (at about $10^{-35}$ s). After this time, the forces become distinct in almost all interactions—they are no longer unified or symmetric. This transition from GUT to electroweak is an example of **spontaneous symmetry breaking**, in which conditions spontaneously evolved to a point where the forces were no longer unified, breaking that symmetry. This is analogous to a phase transition in the universe, and a clever proposal by American physicist Alan Guth in the early 1980s ties it to the smoothness of the CMBR. Guth proposed that spontaneous symmetry breaking (like a phase transition during cooling of normal matter) released an immense amount of energy that caused the universe to expand extremely rapidly for the brief time from $10^{-35}$ s to about $10^{-32}$ s. This expansion may have been by an incredible factor of $10^{50}$ or more in the size of the universe and is thus called the **inflationary scenario**. One result of this inflation is that it would stretch the wrinkles in the universe nearly flat, leaving an extremely smooth CMBR. While speculative, there is as yet no other plausible explanation for the smoothness of the CMBR. Unless the CMBR is not really cosmic but local in origin, the distances between regions of similar temperatures are too great for any coordination to have caused them, since any coordination mechanism must travel at the speed of light. Again, particle physics and cosmology are intimately entwined. There is little hope that we may be able to test the inflationary scenario directly, since it occurs at energies near $10^{14}$ GeV, vastly greater than the limits of modern accelerators. But the idea is so attractive that it is incorporated into most cosmological theories.

Characteristics of the present universe may help us determine the validity of this intriguing idea. Additionally, the recent indications that the universe's expansion rate may be increasing (see [Dark Matter and Closure](#)) could even imply that we are in another inflationary epoch.

It is important to note that, if conditions such as those found in the early universe could be created in the laboratory, we would see the unification of forces directly today. The forces have not changed in time, but the average energy and separation of particles in the universe have. As discussed in [The Four Basic Forces](#), the four basic forces in nature are distinct under most circumstances found today. The early universe and its remnants provide evidence from times when they were unified under most circumstances.

### 34.2 General Relativity and Quantum Gravity

When we talk of black holes or the unification of forces, we are actually discussing aspects of general relativity and quantum gravity. We know from [Special Relativity](#) that relativity is the study of how different observers measure the same event, particularly if they move relative to one another. Einstein's theory of **general relativity** describes all types of relative motion including accelerated motion and the effects of gravity. General relativity encompasses special relativity and classical relativity in situations where acceleration is zero and relative velocity is small compared with the speed of light. Many aspects of general relativity have been verified experimentally, some of which are better than science fiction in that they are bizarre but true. **Quantum gravity** is the theory that deals with particle exchange of gravitons as the mechanism for the force, and with extreme conditions where quantum mechanics and general relativity must both be used. A good theory of quantum gravity does not yet exist, but one will be needed to understand how all four forces may be unified. If we are successful, the theory of quantum gravity will encompass all others, from classical physics to relativity to quantum mechanics—truly a Theory of Everything (TOE).

**General Relativity**

Einstein first considered the case of no observer acceleration when he developed the revolutionary special theory of relativity, publishing his first work on it in 1905. By 1916, he had laid the foundation of general relativity, again almost on his own. Much of what Einstein did to develop his ideas was to mentally analyze certain carefully and clearly defined situations—doing this is to perform a thought experiment. Figure 34.10 illustrates a thought experiment like the ones that convinced Einstein that light must fall in a gravitational field. Think about what a person feels in an elevator. In an elevator, a person feels the effects of gravity. The feet of a person are pressed against the floor, and objects released from hand fall with identical accelerations. In fact, it is not possible, without looking outside, to know what is happening—acceleration upward or gravity. This led Einstein to correctly postulate that acceleration and gravity will produce identical effects in all situations. So, if acceleration affects light, then gravity will, too. Figure 34.10 shows the effect of acceleration on a beam of light shone horizontally at one wall. Since the accelerated elevator moves up during the time light travels across the elevator, the beam of light strikes low, seeming to the person to bend down. (Normally a tiny effect, since the speed of light is so great.) The same effect must occur due to gravity, Einstein reasoned, since there is no way to tell the effects of gravity acting downward from acceleration of the elevator upward. Thus gravity affects the path of light, even though we think of gravity as acting between masses and photons are massless.
Einstein’s theory of general relativity got its first verification in 1919 when starlight passing near the Sun was observed during a solar eclipse. (See Figure 34.11.) During an eclipse, the sky is darkened and we can briefly see stars. Those in a line of sight nearest the Sun should have a shift in their apparent positions. Not only was this shift observed, but it agreed with Einstein’s predictions well within experimental uncertainties. This discovery created a scientific and public sensation. Einstein was now a folk hero as well as a very great scientist. The bending of light by matter is equivalent to a bending of space itself, with light following the curve. This is another radical change in our concept of space and time. It is also another connection that any particle with mass or energy (massless photons) is affected by gravity.

There are several current forefront efforts related to general relativity. One is the observation and analysis of gravitational lensing of light. Another is analysis of the definitive proof of the existence of black holes. Direct observation of gravitational waves or moving wrinkles in space is being searched for. Theoretical efforts are also being aimed at the possibility of time travel and wormholes into other parts of space due to black holes.

**Gravitational lensing** As you can see in Figure 34.11, light is bent toward a mass, producing an effect much like a converging lens (large masses are needed to produce observable effects). On a galactic scale, the light from a distant galaxy could be “lensed” into several images when passing close by another galaxy on its way to Earth. Einstein predicted this effect, but he considered it unlikely that we would ever observe it. A number of cases of this effect have now been observed; one is shown in Figure 34.12. This effect is a much larger scale verification of general relativity. But such gravitational lensing is also useful in verifying that the red shift is proportional to distance. The red shift of the intervening galaxy is always less than that of the one being lensed, and each image of the lensed galaxy has the same red shift. This verification supplies more evidence that red shift is proportional to distance. Confidence that the multiple images are not different objects is bolstered by the observations that if one image varies in brightness over time, the others also vary in the same manner.

**Figure 34.11** This schematic shows how light passing near a massive body like the Sun is curved toward it. The light that reaches the Earth then seems to be coming from different locations than the known positions of the originating stars. Not only was this effect observed, the amount of bending was precisely what Einstein predicted in his general theory of relativity.
Black holes are objects having such large gravitational fields that things can fall in, but nothing, not even light, can escape. Bodies, like the Earth or the Sun, have what is called an escape velocity. If an object moves straight up from the body, starting at the escape velocity, it will just be able to escape the gravity of the body. The greater the acceleration of gravity on the body, the greater is the escape velocity. As long ago as the late 1700s, it was proposed that if the escape velocity is greater than the speed of light, then light cannot escape. Simon Laplace (1749–1827), the French astronomer and mathematician, even incorporated this idea of a dark star into his writings. But the idea was dropped after Young’s double slit experiment showed light to be a wave. For some time, light was thought not to have particle characteristics and, thus, could not be acted upon by gravity. The idea of a black hole was very quickly reincarnated in 1916 after Einstein’s theory of general relativity was published. It is now thought that black holes can form in the supernova collapse of a massive star, forming an object perhaps 10 km across and having a mass greater than that of our Sun. It is interesting that several prominent physicists who worked on the concept, including Einstein, firmly believed that nature would find a way to prohibit such objects.

Black holes are difficult to observe directly, because they are small and no light comes directly from them. In fact, no light comes from inside the event horizon, which is defined to be at a distance from the object at which the escape velocity is exactly the speed of light. The radius of the event horizon is known as the Schwarzschild radius \( R_S \) and is given by

\[
R_S = \frac{2GM}{c^2},
\]

where \( G \) is the universal gravitational constant, \( M \) is the mass of the body, and \( c \) is the speed of light. The event horizon is the edge of the black hole and \( R_S \) is its radius (that is, the size of a black hole is twice \( R_S \)). Since \( G \) is small and \( c^2 \) is large, you can see that black holes are extremely small, only a few kilometers for masses a little greater than the Sun’s. The object itself is inside the event horizon.

Physics near a black hole is fascinating. Gravity increases so rapidly that, as you approach a black hole, the tidal effects tear matter apart, with matter closer to the hole being pulled in with much more force than that only slightly farther away. This can pull a companion star apart and heat inflowing gases to the point of producing X rays. (See Figure 34.13.) We have observed X rays from certain binary star systems that are consistent with such a picture. This is not quite proof of black holes, because the X rays could also be caused by matter falling onto a neutron star. These objects were first discovered in 1967 by the British astrophysicists, Jocelyn Bell and Anthony Hewish. Neutron stars are literally a star composed of neutrons. They are formed by the collapse of a star’s core in a supernova, during which electrons and protons are forced together to form neutrons (the reverse of neutron \( \beta \) decay). Neutron stars are slightly larger than a black hole of the same mass and will not collapse further because of resistance by the strong force. However, neutron stars cannot have a mass greater than about eight solar masses or they must collapse to a black hole. With recent improvements in our ability to resolve small details, such as with the orbiting Chandra X-ray Observatory, it has become possible to measure the masses of X-ray-emitting objects by observing the motion of companion stars and other matter in their vicinity. What has emerged is a plethora of X-ray-emitting objects too massive to be neutron stars. This evidence is considered conclusive and the existence of black holes is widely accepted. These black holes are concentrated near galactic centers.

We also have evidence that supermassive black holes may exist at the cores of many galaxies, including the Milky Way. Such a black hole might have a mass millions or even billions of times that of the Sun, and it would probably have formed when matter first coalesced into a galaxy billions of years ago. Supporting this is the fact that very distant galaxies are more likely to have abnormally energetic cores. Some of the moderately distant galaxies, and hence among the younger, are known as quasars and emit as much or more energy than a normal galaxy but from a region less than a light year across. Quasar energy outputs may vary in times less than a year, so that the energy-emitting region must be less than a light year across. The best explanation of quasars is that they are young galaxies with a supermassive black hole forming at their core, and that they become less energetic over billions of years. In closer superactive galaxies, we observe tremendous amounts of energy being emitted from very small regions of space, consistent with stars falling into a black hole at the rate of one or more a month. The Hubble Space Telescope (1994) observed an accretion disk in the galaxy M87 rotating rapidly around a region of extreme energy emission. (See Figure 34.13.) A jet of material being ejected perpendicular to the plane of rotation gives further evidence of a supermassive black hole as the engine.

\[\text{Figure 34.12} \] (a) Light from a distant galaxy can travel different paths to the Earth because it is bent around an intermediary galaxy by gravity. This produces several images of the more distant galaxy. (b) The images around the central galaxy are produced by gravitational lensing. Each image has the same spectrum and a larger red shift than the intermediary. (credit: NASA, ESA, and STScI)
A black hole is shown pulling matter away from a companion star, forming a superheated accretion disk where X rays are emitted before the matter disappears forever into the hole. The in-fall energy also ejects some material, forming the two vertical spikes. (See also the photograph in Introduction to Frontiers of Physics.) There are several X-ray-emitting objects in space that are consistent with this picture and are likely to be black holes.

**Gravitational waves** If a massive object distorts the space around it, like the foot of a water bug on the surface of a pond, then movement of the massive object should create waves in space like those on a pond. Gravitational waves are mass-created distortions in space that propagate at the speed of light and are predicted by general relativity. Since gravity is by far the weakest force, extreme conditions are needed to generate significant gravitational waves. Gravity near binary neutron star systems is so great that significant gravitational wave energy is radiated as the two neutron stars orbit one another. American astronomers, Joseph Taylor and Russell Hulse, measured changes in the orbit of such a binary neutron star system. They found its orbit to change precisely as predicted by general relativity, a strong indication of gravitational waves, and were awarded the 1993 Nobel Prize. But direct detection of gravitational waves on Earth would be conclusive. For many years, various attempts have been made to detect gravitational waves by observing vibrations induced in matter distorted by these waves. American physicist Joseph Weber pioneered this field in the 1960s, but no conclusive events have been observed. (No gravity wave detectors were in operation at the time of the 1987A supernova, unfortunately.) There are now several ambitious systems of gravitational wave detectors in use around the world. These include the LIGO (Laser Interferometer Gravitational Wave Observatory) system with two laser interferometer detectors, one in the state of Washington and another in Louisiana (See Figure 34.15) and the VIRGO (Variability of Irradiance and Gravitational Oscillations) facility in Italy with a single detector.

**Quantum Gravity**

Black holes radiate Quantum gravity is important in those situations where gravity is so extremely strong that it has effects on the quantum scale, where the other forces are ordinarily much stronger. The early universe was such a place, but black holes are another. The first significant connection between gravity and quantum effects was made by the Russian physicist Yakov Zel’dovich in 1971, and other significant advances followed from the British physicist Stephen Hawking. (See Figure 34.16.) These two showed that black holes could radiate away energy by quantum effects just outside the event horizon (nothing can escape from inside the event horizon). Black holes are, thus, expected to radiate energy and shrink to nothing, although extremely slowly for most black holes. The mechanism is the creation of a particle-antiparticle pair from energy in the extremely strong gravitational field near the event horizon. One member of the pair falls into the hole and the other escapes, conserving momentum. (See Figure 34.17.) When a black hole loses energy and, hence, rest mass, its event horizon shrinks, creating an even greater gravitational field. This increases the rate of pair production so that the process grows exponentially until the black hole is nuclear in size. A final burst of particles and $\gamma$ rays ensues. This is an extremely slow process for black holes about the mass of the Sun (produced by supernovas) or larger ones (like those thought to be at galactic centers), taking on the order of $10^{67}$ years or longer! Smaller black holes would evaporate faster, but they are only speculated to exist as remnants of the Big Bang. Searches for characteristic $\gamma$-ray bursts have produced events attributable to more mundane objects like neutron stars accreting matter.

**Figure 34.13** A black hole is shown pulling matter away from a companion star, forming a superheated accretion disk where X rays are emitted before the matter disappears forever into the hole. The in-fall energy also ejects some material, forming the two vertical spikes. (See also the photograph in Introduction to Frontiers of Physics.) There are several X-ray-emitting objects in space that are consistent with this picture and are likely to be black holes.

**Figure 34.14** This Hubble Space Telescope photograph shows the extremely energetic core of the NGC 4261 galaxy. With the superior resolution of the orbiting telescope, it has been possible to observe the rotation of an accretion disk around the energy-producing object as well as to map jets of material being ejected from the object. A supermassive black hole is consistent with these observations, but other possibilities are not quite eliminated. (credit: NASA and ESA)
Gravitational waves will cause extremely small vibrations in a mass in this detector, which will be detected by laser interferometer techniques. Such detection in coincidence with other detectors and with astronomical events, such as supernovas, would provide direct evidence of gravitational waves. (credit: Tobin Fricke)

Stephen Hawking (b. 1942) has made many contributions to the theory of quantum gravity. Hawking is a long-time survivor of ALS and has produced popular books on general relativity, cosmology, and quantum gravity. (credit: Lwp Kommunikáció)

Gravity and quantum mechanics come into play when a black hole creates a particle-antiparticle pair from the energy in its gravitational field. One member of the pair falls into the hole while the other escapes, removing energy and shrinking the black hole. The search is on for the characteristic energy.

Wormholes and time travel

The subject of time travel captures the imagination. Theoretical physicists, such as the American Kip Thorne, have treated the subject seriously, looking into the possibility that falling into a black hole could result in popping up in another time and place—a trip through a so-called wormhole. Time travel and wormholes appear in innumerable science fiction dramatizations, but the consensus is that time travel is not possible in theory. While still debated, it appears that quantum gravity effects inside a black hole prevent time travel due to the creation of particle pairs. Direct evidence is elusive.

The shortest time

Theoretical studies indicate that, at extremely high energies and correspondingly early in the universe, quantum fluctuations may make time intervals meaningful only down to some finite time limit. Early work indicated that this might be the case for times as long as $10^{-43}$ s, the time at which all forces were unified. If so, then it would be meaningless to consider the universe at times earlier than this. Subsequent studies indicate that the crucial time may be as short as $10^{-95}$ s. But the point remains—quantum gravity seems to imply that there is no such thing as a vanishingly short time. Time may, in fact, be grainy with no meaning to time intervals shorter than some tiny but finite size.
The future of quantum gravity

Not only is quantum gravity in its infancy, no one knows how to get started on a theory of gravitons and unification of forces. The energies at which TOE should be valid may be so high (at least $10^{19}$ GeV) and the necessary particle separation so small (less than $10^{-35}$ m) that only indirect evidence can provide clues. For some time, the common lament of theoretical physicists was one so familiar to struggling students—how do you even get started? But Hawking and others have made a start, and the approach many theorists have taken is called Superstring theory, the topic of the Superstrings.

### 34.3 Superstrings

Introduced earlier in GUTS: The Unification of Forces, Superstring theory is an attempt to unify gravity with the other three forces and, thus, must contain quantum gravity. The main tenet of Superstring theory is that fundamental particles, including the graviton that carries the gravitational force, act like one-dimensional vibrating strings. Since gravity affects the time and space in which all else exists, Superstring theory is an attempt at a Theory of Everything (TOE). Each independent quantum number is thought of as a separate dimension in some super space (analogous to the fact that the familiar dimensions of space are independent of one another) and is represented by a different type of Superstring. As the universe evolved after the Big Bang and forces became distinct (spontaneous symmetry breaking), some of the dimensions of superspace are imagined to have curled up and become unnoticed.

Forces are expected to be unified only at extremely high energies and at particle separations on the order of $10^{-35}$ m. This could mean that Superstrings must have dimensions or wavelengths of this size or smaller. Just as quantum gravity may imply that there are no time intervals shorter than some finite value, it also implies that there may be no sizes smaller than some tiny but finite value. That may be about $10^{-35}$ m. If so, and if Superstring theory can explain all it strives to, then the structures of Superstrings are at the lower limit of the smallest possible size and can have no further substructure. This would be the ultimate answer to the question the ancient Greeks considered. There is a finite lower limit to space.

Not only is Superstring theory in its infancy, it deals with dimensions about 17 orders of magnitude smaller than the $10^{-18}$ m details that we have been able to observe directly. It is thus relatively unconstrained by experiment, and there are a host of theoretical possibilities to choose from. This has led theorists to make choices subjectively (as always) on what is the most elegant theory, with less hope than usual that experiment will guide them. It has also led to speculation of alternate universes, with their Big Bangs creating each new universe with a random set of rules. These speculations may not be tested even in principle, since an alternate universe is by definition unattainable. It is something like exploring a self-consistent field of mathematics, with its axioms and rules of logic that are not consistent with nature. Such endeavors have often given insight to mathematicians and scientists alike and occasionally have been directly related to the description of new discoveries.

### 34.4 Dark Matter and Closure

One of the most exciting problems in physics today is the fact that there is far more matter in the universe than we can see. The motion of stars in galaxies and the motion of galaxies in clusters imply that there is about 10 times as much mass as in the luminous objects we can see. The indirectly observed non-luminous matter is called dark matter. Why is dark matter a problem? For one thing, we do not know what it is. It may well be 90% of all matter in the universe, yet there is a possibility that it is of a completely unknown form—a stunning discovery if verified. Dark matter has implications for particle physics. It may be possible that neutrinos actually have small masses or that there are completely unknown types of particles. Dark matter also has implications for cosmology, since there may be enough dark matter to stop the expansion of the universe. That is another problem related to dark matter—we do not know how much there is. We keep finding evidence for more matter in the universe, and we have an idea of how much it would take to eventually stop the expansion of the universe, but whether there is enough is still unknown.

#### Evidence

The first clues that there is more matter than meets the eye came from the Swiss-born American astronomer Fritz Zwicky in the 1930s; some initial work was also done by the American astronomer Vera Rubin. Zwicky measured the velocities of stars orbiting the galaxy, using the relativistic Doppler shift of their spectra (see Figure 34.18(a)). He found that velocity varied with distance from the center of the galaxy, as graphed in Figure 34.18(b). If the mass of the galaxy was concentrated in its center, as are its luminous stars, the velocities should decrease as the square root of the distance from the center. Instead, the velocity curve is almost flat, implying that there is a tremendous amount of matter in the galactic halo. Although not immediately recognized for its significance, such measurements have now been made for many galaxies, with similar results. Further, studies of galactic clusters have also indicated that galaxies have a mass distribution greater than that obtained from their brightness (proportional to the number of stars), which also extends into large halos surrounding the luminous parts of galaxies. Observations of other EM wavelengths, such as radio waves and X rays, have similarly confirmed the existence of dark matter. Take, for example, X rays in the relatively dark space between galaxies, which indicates the presence of previously unobserved hot, ionized gas (see Figure 34.18(c)).

#### Theoretical Yearnings for Closure

Is the universe open or closed? That is, will the universe expand forever or will it stop, perhaps to contract? This, until recently, was a question of whether there is enough gravitation to stop the expansion of the universe. In the past few years, it has become a question of the combination of gravitation and what is called the cosmological constant. The cosmological constant was invented by Einstein to prohibit the expansion or contraction of the universe. Einstein later discarded that an illogical possibility. The cosmological constant was resurrected after Hubble discovered the expansion, but has been re-invoked in recent years.

Gravitational attraction between galaxies is slowing the expansion of the universe, but the amount of slowing down is not known directly. In fact, the cosmological constant can counteract gravity’s effect. As recent measurements indicate, the universe is expanding faster now than in the past—perhaps a “modern inflationary era” in which the dark energy is thought to be causing the expansion of the present-day universe to accelerate. If the expansion rate were affected by gravity alone, we should be able to see that the expansion rate between distant galaxies was once greater than it is now. However, measurements show it was less than now. We can, however, calculate the amount of slowing based on the average density of matter we observe directly. Here we have a definite answer—there is far less visible matter than needed to stop expansion. The critical density $\rho_c$ is defined to be the density needed to just halt universal expansion in a universe with no cosmological constant. It is estimated to be about

$$\rho_c \approx 10^{-26} \text{ kg/m}^3.$$  

(34.3)
However, this estimate of $\rho_c$ is only good to about a factor of two, due to uncertainties in the expansion rate of the universe. The critical density is equivalent to an average of only a few nucleons per cubic meter, remarkably small and indicative of how truly empty intergalactic space is. Luminous matter seems to account for roughly 0.5% to 2% of the critical density, far less than that needed for closure. Taking into account the amount of dark matter we detect indirectly and all other types of indirectly observed normal matter, there is only 10% to 40% of what is needed for closure. If we are able to refine the measurements of expansion rates now and in the past, we will have our answer regarding the curvature of space and we will determine a value for the cosmological constant to justify this observation. Finally, the most recent measurements of the CMBR have implications for the cosmological constant, so it is not simply a device concocted for a single purpose.

After the recent experimental discovery of the cosmological constant, most researchers feel that the universe should be just barely open. Since matter can be thought to curve the space around it, we call an open universe negatively curved. This means that you can in principle travel an unlimited distance in any direction. A universe that is closed is called positively curved. This means that if you travel far enough in any direction, you will return to your starting point, analogous to circumnavigating the Earth. In between these two is a flat (zero curvature) universe. The recent discovery of the cosmological constant has shown the universe is very close to flat, and will expand forever. Why do theorists feel the universe is flat? Flatness is a part of the inflationary scenario that helps explain the flatness of the microwave background. In fact, since general relativity implies that matter creates the space in which it exists, there is a special symmetry to a flat universe.

**Figure 34.18** Evidence for dark matter: (a) We can measure the velocities of stars relative to their galaxies by observing the Doppler shift in emitted light, usually using the hydrogen spectrum. These measurements indicate the rotation of a spiral galaxy. (b) A graph of velocity versus distance from the galactic center shows that the velocity does not decrease as it would if the matter were concentrated in luminous stars. The flatness of the curve implies a massive galactic halo of dark matter extending beyond the visible stars. (c) This is a computer-generated image of X rays from a galactic cluster. The X rays indicate the presence of otherwise unseen hot clouds of ionized gas in the regions of space previously considered more empty. (credit: NASA, ESA, CXC, M. Bradac (University of California, Santa Barbara), and S. Allen (Stanford University))
What Is the Dark Matter We See Indirectly?

There is no doubt that dark matter exists, but its form and the amount in existence are two facts that are still being studied vigorously. As always, we seek to explain new observations in terms of known principles. However, as more discoveries are made, it is becoming more and more difficult to explain dark matter as a known type of matter.

One of the possibilities for normal matter is being explored using the Hubble Space Telescope and employing the lensing effect of gravity on light (see Figure 34.19). Stars glow because of nuclear fusion in them, but planets are visible primarily by reflected light. Jupiter, for example, is too small to ignite fusion in its core and become a star, but we can see sunlight reflected from it, since we are relatively close. If Jupiter orbited another star, we would not be able to see it directly. The question is open as to how many planets or other bodies smaller than about 1/1000 the mass of the Sun are there. If such bodies pass between us and a star, they will not block the star’s light, being too small, but they will form a gravitational lens, as discussed in General Relativity and Quantum Gravity.

In a process called microlensing, light from the star is focused and the star appears to brighten in a characteristic manner. Searches for dark matter in this form are particularly interested in galactic halos because of the huge amount of mass that seems to be there. Such microlensing objects are thus called massive compact halo objects, or MACHOs. To date, a few MACHOs have been observed, but not predominantly in galactic halos, nor in the numbers needed to explain dark matter.

MACHOs are among the most conventional of unseen objects proposed to explain dark matter. Others being actively pursued are red dwarfs, which are small dim stars, but too few have been seen so far, even with the Hubble Telescope, to be of significance. Old remnants of stars called white dwarfs are also under consideration, since they contain about a solar mass, but are small as the Earth and may dim to the point that we ordinarily do not observe them. While white dwarfs are known, old dim ones are not. Yet another possibility is the existence of large numbers of smaller than stellar mass black holes left from the Big Bang—here evidence is entirely absent.

There is a very real possibility that dark matter is composed of the known neutrinos, which may have small, but finite, masses. As discussed earlier, neutrinos are thought to be massless, but we only have upper limits on their masses, rather than knowing they are exactly zero. So far, these upper limits come from difficult measurements of total energy emitted in the decays and reactions in which neutrinos are involved. There is an amusing possibility of proving that neutrinos have mass in a completely different way.

We have noted in Particles, Patterns, and Conservation Laws that there are three flavors of neutrinos (ν_e, ν_µ, and ν_τ) and that the weak interaction could change quark flavor. It should also change neutrino flavor—that is, any type of neutrino could change spontaneously into any other, a process called neutrino oscillations. However, this can occur only if neutrinos have a mass. Why? Crudely, because if neutrinos are massless, they must travel at the speed of light and time will not pass for them, so that they cannot change without an interaction. In 1999, results began to be published containing convincing evidence that neutrino oscillations do occur. Using the Super-Kamiokande detector in Japan, the oscillations have been observed and are being verified and further explored at present at the same facility and others.

Neutrino oscillations may also explain the low number of observed solar neutrinos. Detectors for observing solar neutrinos are specifically designed to detect electron neutrinos ν_e produced in huge numbers by fusion in the Sun. A large fraction of electron neutrinos ν_e may be changing flavor to muon neutrinos ν_µ on their way out of the Sun, possibly enhanced by specific interactions, reducing the flux of electron neutrinos to observed levels.

There is also a discrepancy in observations of neutrinos produced in cosmic ray showers. While these showers of radiation produced by extremely energetic cosmic rays should contain twice as many ν_µ s as ν_e s, their numbers are nearly equal. This may be explained by neutrino oscillations from muon flavor to electron flavor. Massive neutrinos are a particularly appealing possibility for explaining dark matter, since their existence is consistent with a large body of known information and explains more than dark matter. The question is not settled at this writing.

The most radical proposal to explain dark matter is that it consists of previously unknown leptons (sometimes obtusely referred to as non-baryonic matter). These are called weakly interacting massive particles, or WIMPs, and would also be chargeless, thus interacting negligibly with normal matter, except through gravitation. One proposed group of WIMPs would have masses several orders of magnitude greater than nucleons and are sometimes called neutralinos. Others are called axions and would have masses about 10^{-10} that of an electron mass. Both neutralinos and axions would be gravitationally attached to galaxies, but because they are chargeless and only feel the weak force, they would be in a halo rather than interact and coalesce into spirals, and so on, like normal matter (see Figure 34.20).

Figure 34.19 The Hubble Space Telescope is producing exciting data with its corrected optics and with the absence of atmospheric distortion. It has observed some MACHOs, disks of material around stars thought to precede planet formation, black hole candidates, and collisions of comets with Jupiter. (credit: NASA (crew of STS-125))

Figure 34.20
Some particle theorists have built WIMPs into their unified force theories and into the inflationary scenario of the evolution of the universe so popular today. These particles would have been produced in just the correct numbers to make the universe flat, shortly after the Big Bang. The proposal is radical in the sense that it invokes entirely new forms of matter, in fact two entirely new forms, in order to explain dark matter and other phenomena. WIMPs have the extra burden of automatically being very difficult to observe directly. This is somewhat analogous to quark confinement, which guarantees that quarks are there, but they can never be seen directly. One of the primary goals of the LHC at CERN, however, is to produce and detect WIMPs. At any rate, before WIMPs are accepted as the best explanation, all other possibilities utilizing known phenomena will have to be shown inferior. Should that occur, we will be in the unanticipated position of admitting that, to date, all we know is only 10% of what exists. A far cry from the days when people firmly believed themselves to be not only the center of the universe, but also the reason for its existence.

34.5 Complexity and Chaos

Much of what impresses us about physics is related to the underlying connections and basic simplicity of the laws we have discovered. The language of physics is precise and well defined because many basic systems we study are simple enough that we can perform controlled experiments and discover unambiguous relationships. Our most spectacular successes, such as the prediction of previously unobserved particles, come from the simple underlying patterns we have been able to recognize. But there are systems of interest to physicists that are inherently complex. The simple laws of physics apply, of course, but complex systems may reveal patterns that simple systems do not. The emerging field of complexity is devoted to the study of complex systems, including those outside the traditional bounds of physics. Of particular interest is the ability of complex systems to adapt and evolve.

What are some examples of complex adaptive systems? One is the primordial ocean. When the oceans first formed, they were a random mix of elements and compounds that obeyed the laws of physics and chemistry. In a relatively short geological time (about 500 million years), life had emerged. Laboratory simulations indicate that the emergence of life was far too fast to have come from random combinations of compounds, even if driven by lightning and heat. There must be an underlying ability of the complex system to organize itself, resulting in the self-replication we recognize as life. Living entities, even at the unicellular level, are highly organized and systematic. Systems of living organisms are themselves complex adaptive systems. The grandest of these evolved into the biological system we have today, leaving traces in the geological record of steps taken along the way.

Complexity as a discipline examines complex systems, how they adapt and evolve, looking for similarities with other complex adaptive systems. Can, for example, parallels be drawn between biological evolution and the evolution of economic systems? Economic systems do emerge quickly, they show tendencies for self-organization, they are complex (in the number and types of transactions), and they adapt and evolve. Biological systems do all the same types of things. There are other examples of complex adaptive systems being studied for fundamental similarities. Cultures show signs of adaptation and evolution. The comparison of different cultural evolutions may bear fruit as well as comparisons to biological evolution. Science also is a complex system of human interactions, like culture and economics, that adapts to new information and political pressure, and evolves, usually becoming more organized rather than less. Those who study creative thinking also see parallels with complex systems. Humans sometimes organize almost random pieces of information, often subconsciously while doing other things, and come up with brilliant creative insights. The development of language is another complex adaptive system that may show similar tendencies. Artificial intelligence is an overt attempt to devise an adaptive system that will self-organize and evolve in the same manner as an intelligent living being learns. These are a few of the broad range of topics being studied by those who investigate complexity. There are now institutes, journals, and meetings, as well as popularizations of the emerging topic of complexity.

In traditional physics, the discipline of complexity may yield insights in certain areas. Thermodynamics treats systems on the average, while statistical mechanics deals in some detail with complex systems of atoms and molecules in random thermal motion. Yet there is organization, adaptation, and evolution in these complex systems. Non-equilibrium phenomena, such as heat transfer and phase changes, are characteristically complex in detail, and new approaches to them may evolve from complexity as a discipline. Crystal growth is another example of self-organization spontaneously emerging in a complex system. Alloys are also inherently complex mixtures that show certain simple characteristics implying some self-organization. The organization of iron atoms into magnetic domains as they cool is another. Perhaps insights into these difficult areas will emerge from complexity. But at the minimum, the discipline of complexity is another example of human effort to understand and organize the universe around us, partly rooted in the discipline of physics.

A predecessor to complexity is the topic of chaos, which has been widely publicized and has become a discipline of its own. It is also based partly in physics and treats broad classes of phenomena from many disciplines. Chaos is a word used to describe systems whose outcomes are extremely sensitive to initial conditions. The orbit of the planet Pluto, for example, may be chaotic in that it can change tremendously due to small interactions with other planets. This makes its long-term behavior impossible to predict with precision, just as we cannot tell precisely where a decaying Earth satellite will land or how many pieces it will break into. But the discipline of chaos has found ways to deal with such systems and has been applied to apparently unrelated systems. For example, the heartbeat of people with certain types of potentially lethal arrhythmias seems to be chaotic, and this knowledge may allow more sophisticated monitoring and recognition of the need for intervention.
Chaos is related to complexity. Some chaotic systems are also inherently complex; for example, vortices in a fluid as opposed to a double pendulum. Both are chaotic and not predictable in the same sense as other systems. But there can be organization in chaos and it can also be quantified. Examples of chaotic systems are beautiful fractal patterns such as in Figure 34.21. Some chaotic systems exhibit self-organization, a type of stable chaos. The orbits of the planets in our solar system, for example, may be chaotic (we are not certain yet). But they are definitely organized and systematic, with a simple formula describing the orbital radii of the first eight planets and the asteroid belt. Large-scale vortices in Jupiter’s atmosphere are chaotic, but the Great Red Spot is a stable self-organization of rotational energy. (See Figure 34.22.) The Great Red Spot has been in existence for at least 400 years and is a complex self-adaptive system.

The emerging field of complexity, like the now almost traditional field of chaos, is partly rooted in physics. Both attempt to see similar systematics in a very broad range of phenomena and, hence, generate a better understanding of them. Time will tell what impact these fields have on more traditional areas of physics as well as on the other disciplines they relate to.

Figure 34.21 This image is related to the Mandelbrot set, a complex mathematical form that is chaotic. The patterns are infinitely fine as you look closer and closer, and they indicate order in the presence of chaos. (credit: Gilberto Santa Rosa)

Figure 34.22 The Great Red Spot on Jupiter is an example of self-organization in a complex and chaotic system. Smaller vortices in Jupiter’s atmosphere behave chaotically, but the triple-Earth-size spot is self-organized and stable for at least hundreds of years. (credit: NASA)

### 34.6 High-temperature Superconductors

Superconductors are materials with a resistivity of zero. They are familiar to the general public because of their practical applications and have been mentioned at a number of points in the text. Because the resistance of a piece of superconductor is zero, there are no heat losses for currents through them; they are used in magnets needing high currents, such as in MRI machines, and could cut energy losses in power transmission. But most superconductors must be cooled to temperatures only a few kelvin above absolute zero, a costly procedure limiting their practical applications. In the past decade, tremendous advances have been made in producing materials that become superconductors at relatively high temperatures. There is hope that room temperature superconductors may someday be manufactured.

Superconductivity was discovered accidentally in 1911 by the Dutch physicist H. Kamerlingh Onnes (1853–1926) when he used liquid helium to cool mercury. Onnes had been the first person to liquefy helium a few years earlier and was surprised to observe the resistivity of a mediocre conductor like mercury drop to zero at a temperature of 4.2 K. We define the temperature at which and below which a material becomes a superconductor to be its critical temperature, denoted by $T_c$. (See Figure 34.23.) Progress in understanding how and why a material became a superconductor was relatively slow, with the first workable theory coming in 1957. Certain other elements were also found to become superconductors, but all had $T_c$ s less than 10 K, which are expensive to maintain. Although Onnes received a Nobel prize in 1913, it was primarily for his work with liquid helium.

In 1986, a breakthrough was announced—a ceramic compound was found to have an unprecedented $T_c$ of 35 K. It looked as if much higher critical temperatures could be possible, and by early 1988 another ceramic (this of thallium, calcium, barium, copper, and oxygen) had been found to have $T_c = 125$ K (see Figure 34.24.). The economic potential of perfect conductors saving electric energy is immense for $T_c$ s above 77 K, since that is the temperature of liquid nitrogen. Although liquid helium has a boiling point of 4 K and can be used to make materials superconducting, it costs about $5 per liter. Liquid nitrogen boils at 77 K, but only costs about $0.30 per liter. There was general euphoria at the discovery of these complex ceramic superconductors, but this soon subsided with the sobering difficulty of forming them into usable wires. The first commercial use of a high temperature superconductor is in an electronic filter for cellular phones. High-temperature superconductors are used in experimental apparatus, and they are actively being researched, particularly in thin film applications.
High temperature superconductors have verifiable \( T_c \)s greater than 125 K, well above the easily achieved 77-K temperature of liquid nitrogen. The search is on for even higher \( T_c \) superconductors, many of complex and exotic copper oxide ceramics, sometimes including strontium, mercury, or yttrium as well as barium, calcium, and other elements. Room temperature (about 293 K) would be ideal, but any temperature close to room temperature is relatively cheap to produce and maintain. There are persistent reports of \( T_c \)s over 200 K and some in the vicinity of 270 K.

Unfortunately, these observations are not routinely reproducible, with samples losing their superconducting nature once heated and recooled (cycled) a few times (see Figure 34.25). They are now called USOs or unidentified superconducting objects, out of frustration and the refusal of some samples to show high \( T_c \) even though produced in the same manner as others. Reproducibility is crucial to discovery, and researchers are justifiably reluctant to claim the breakthrough they all seek. Time will tell whether USOs are real or an experimental quirk.

The theory of ordinary superconductors is difficult, involving quantum effects for widely separated electrons traveling through a material. Electrons couple in a manner that allows them to get through the material without losing energy to it, making it a superconductor. High-\( T_c \) superconductors are more difficult to understand theoretically, but theorists seem to be closing in on a workable theory. The difficulty of understanding how electrons can sneak through materials without losing energy in collisions is even greater at higher temperatures, where vibrating atoms should get in the way. Discoverers of high \( T_c \) may feel something analogous to what a politician once said upon an unexpected election victory—“I wonder what we did right?”
A  ATOMIC MASSES


## Table A1: Atomic Masses

<table>
<thead>
<tr>
<th>Atomic Number, Z</th>
<th>Name</th>
<th>Atomic Number, A</th>
<th>Symbol</th>
<th>Atomic Mass (u)</th>
<th>Percent Abundance or Decay Mode</th>
<th>Half-life, ( t_{1/2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>neutron</td>
<td>1</td>
<td>( n )</td>
<td>1.008 665</td>
<td>( \beta^- )</td>
<td>10.37 min</td>
</tr>
<tr>
<td>1</td>
<td>Hydrogen</td>
<td>1</td>
<td>( ^1\text{H} )</td>
<td>1.007 825</td>
<td>99.985%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deuterium</td>
<td>2</td>
<td>( ^2\text{H or D} )</td>
<td>2.014 102</td>
<td>0.015%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tritium</td>
<td>3</td>
<td>( ^3\text{H or T} )</td>
<td>3.016 050</td>
<td>( \beta^- )</td>
<td>12.33 y</td>
</tr>
<tr>
<td>2</td>
<td>Helium</td>
<td>3</td>
<td>( ^3\text{He} )</td>
<td>3.016 030</td>
<td>1.38\times10^{-4}%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>( ^4\text{He} )</td>
<td>4.002 603</td>
<td>( \approx100% )</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Lithium</td>
<td>6</td>
<td>( ^6\text{Li} )</td>
<td>6.015 121</td>
<td>7.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>( ^7\text{Li} )</td>
<td>7.016 003</td>
<td>92.5%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Beryllium</td>
<td>7</td>
<td>( ^7\text{Be} )</td>
<td>7.016 928</td>
<td>EC</td>
<td>53.29 d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>( ^9\text{Be} )</td>
<td>9.012 182</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Boron</td>
<td>10</td>
<td>( ^{10}\text{B} )</td>
<td>10.012 937</td>
<td>19.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>( ^{11}\text{B} )</td>
<td>11.009 305</td>
<td>80.1%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Carbon</td>
<td>11</td>
<td>( ^{11}\text{C} )</td>
<td>11.011 432</td>
<td>EC, ( \beta^+ )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>( ^{12}\text{C} )</td>
<td>12.000 000</td>
<td>98.90%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>( ^{13}\text{C} )</td>
<td>13.003 355</td>
<td>1.10%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>( ^{14}\text{C} )</td>
<td>14.003 241</td>
<td>( \beta^- )</td>
<td>5730 y</td>
</tr>
<tr>
<td>7</td>
<td>Nitrogen</td>
<td>13</td>
<td>( ^{13}\text{N} )</td>
<td>13.005 738</td>
<td>( \beta^+ )</td>
<td>9.96 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>( ^{14}\text{N} )</td>
<td>14.003 074</td>
<td>99.63%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>( ^{15}\text{N} )</td>
<td>15.000 108</td>
<td>0.37%</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Oxygen</td>
<td>15</td>
<td>( ^{15}\text{O} )</td>
<td>15.003 065</td>
<td>EC, ( \beta^+ )</td>
<td>122 s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>( ^{16}\text{O} )</td>
<td>15.994 915</td>
<td>99.76%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>18</td>
<td>( ^{18}\text{O} )</td>
<td>17.999 160</td>
<td>0.200%</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Fluorine</td>
<td>18</td>
<td>( ^{18}\text{F} )</td>
<td>18.000 937</td>
<td>EC, ( \beta^+ )</td>
<td>1.83 h</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>( ^{19}\text{F} )</td>
<td>18.998 403</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Neon</td>
<td>20</td>
<td>( ^{20}\text{Ne} )</td>
<td>19.992 435</td>
<td>90.51%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>( ^{22}\text{Ne} )</td>
<td>21.991 383</td>
<td>9.22%</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Sodium</td>
<td>22</td>
<td>( ^{22}\text{Na} )</td>
<td>21.994 434</td>
<td>( \beta^+ )</td>
<td>2.602 y</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>( ^{23}\text{Na} )</td>
<td>22.989 767</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>( ^{24}\text{Na} )</td>
<td>23.990 961</td>
<td>( \beta^- )</td>
<td>14.96 h</td>
</tr>
<tr>
<td>12</td>
<td>Magnesium</td>
<td>24</td>
<td>( ^{24}\text{Mg} )</td>
<td>23.985 042</td>
<td>78.99%</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Aluminum</td>
<td>27</td>
<td>( ^{27}\text{Al} )</td>
<td>26.981 539</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Silicon</td>
<td>28</td>
<td>( ^{28}\text{Si} )</td>
<td>27.976 927</td>
<td>92.23%</td>
<td>2.62 h</td>
</tr>
<tr>
<td>Atomic Number, Z</td>
<td>Name</td>
<td>Atomic Mass Number, A</td>
<td>Symbol</td>
<td>Atomic Mass (u)</td>
<td>Percent Abundance or Decay Mode</td>
<td>Half-life, t₁/₂</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
<td>-----------------------</td>
<td>--------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>15</td>
<td>Phosphorus</td>
<td>31</td>
<td>Si</td>
<td>30.975 362</td>
<td>β⁻</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
<td>P</td>
<td>30.973 762</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>P</td>
<td>31.973 907</td>
<td>β⁻</td>
<td>14.28 d</td>
</tr>
<tr>
<td>16</td>
<td>Sulfur</td>
<td>32</td>
<td>S</td>
<td>31.972 070</td>
<td>95.02%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>S</td>
<td>34.969 031</td>
<td>β⁻</td>
<td>87.4 d</td>
</tr>
<tr>
<td>17</td>
<td>Chlorine</td>
<td>35</td>
<td>Cl</td>
<td>34.968 852</td>
<td>75.77%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>37</td>
<td>Cl</td>
<td>36.965 903</td>
<td>24.23%</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Argon</td>
<td>40</td>
<td>Ar</td>
<td>39.962 384</td>
<td>99.60%</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Potassium</td>
<td>39</td>
<td>K</td>
<td>38.963 707</td>
<td>93.26%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>K</td>
<td>39.963 999</td>
<td>0.0117%, EC, β⁻</td>
<td>1.28×10⁹ y</td>
</tr>
<tr>
<td>20</td>
<td>Calcium</td>
<td>40</td>
<td>Ca</td>
<td>39.962 591</td>
<td>96.94%</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Scandium</td>
<td>45</td>
<td>Sc</td>
<td>44.955 910</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Titanium</td>
<td>48</td>
<td>Ti</td>
<td>47.947 947</td>
<td>73.8%</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Vanadium</td>
<td>51</td>
<td>V</td>
<td>50.943 962</td>
<td>99.75%</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Chromium</td>
<td>52</td>
<td>Cr</td>
<td>51.940 509</td>
<td>83.79%</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Manganese</td>
<td>55</td>
<td>Mn</td>
<td>54.938 047</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Iron</td>
<td>56</td>
<td>Fe</td>
<td>55.934 939</td>
<td>91.72%</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Cobalt</td>
<td>59</td>
<td>Co</td>
<td>58.933 198</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>Co</td>
<td>59.933 819</td>
<td>β⁻</td>
<td>5.271 y</td>
</tr>
<tr>
<td>28</td>
<td>Nickel</td>
<td>58</td>
<td>Ni</td>
<td>57.935 346</td>
<td>68.27%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>Ni</td>
<td>59.930 788</td>
<td>26.10%</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Copper</td>
<td>63</td>
<td>Cu</td>
<td>62.939 598</td>
<td>69.17%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>65</td>
<td>Cu</td>
<td>64.927 793</td>
<td>30.83%</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Zinc</td>
<td>64</td>
<td>Zn</td>
<td>63.929 145</td>
<td>48.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>66</td>
<td>Zn</td>
<td>65.926 034</td>
<td>27.9%</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Gallium</td>
<td>69</td>
<td>Ga</td>
<td>68.925 580</td>
<td>60.1%</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Germanium</td>
<td>72</td>
<td>Ge</td>
<td>71.922 079</td>
<td>27.4%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>74</td>
<td>Ge</td>
<td>73.921 177</td>
<td>36.5%</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Arsenic</td>
<td>75</td>
<td>As</td>
<td>74.921 594</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Selenium</td>
<td>80</td>
<td>Se</td>
<td>79.916 520</td>
<td>49.7%</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Bromine</td>
<td>79</td>
<td>Br</td>
<td>78.918 336</td>
<td>50.69%</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Krypton</td>
<td>84</td>
<td>Kr</td>
<td>83.911 507</td>
<td>57.0%</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Rubidium</td>
<td>85</td>
<td>Rb</td>
<td>84.911 794</td>
<td>72.17%</td>
<td></td>
</tr>
<tr>
<td>Atomic Number, Z</td>
<td>Name</td>
<td>Atomic Mass Number, A</td>
<td>Symbol</td>
<td>Atomic Mass (u)</td>
<td>Percent Abundance or Decay Mode</td>
<td>Half-life, $t_{\frac{1}{2}}$</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------------</td>
<td>--------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>38</td>
<td>Strontium</td>
<td>86</td>
<td>$^{86}$Sr</td>
<td>85.909 267</td>
<td>9.86%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>88</td>
<td>$^{88}$Sr</td>
<td>87.905 619</td>
<td>82.58%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>$^{90}$Sr</td>
<td>89.907 738</td>
<td>$\beta^-$</td>
<td>28.8 y</td>
</tr>
<tr>
<td>39</td>
<td>Yttrium</td>
<td>89</td>
<td>$^{89}$Y</td>
<td>88.905 849</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>$^{90}$Y</td>
<td>89.907 152</td>
<td>$\beta^-$</td>
<td>64.1 h</td>
</tr>
<tr>
<td>40</td>
<td>Zirconium</td>
<td>90</td>
<td>$^{90}$Zr</td>
<td>89.904 703</td>
<td>51.45%</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Niobium</td>
<td>93</td>
<td>$^{93}$Nb</td>
<td>92.906 377</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Molybdenum</td>
<td>98</td>
<td>$^{98}$Mo</td>
<td>97.905 406</td>
<td>24.13%</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Technetium</td>
<td>98</td>
<td>$^{98}$Tc</td>
<td>97.907 215</td>
<td>$\beta^-$</td>
<td>4.2×10$^6$ y</td>
</tr>
<tr>
<td>44</td>
<td>Ruthenium</td>
<td>102</td>
<td>$^{102}$Ru</td>
<td>101.904 348</td>
<td>31.6%</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Rhodium</td>
<td>103</td>
<td>$^{103}$Rh</td>
<td>102.905 500</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Palladium</td>
<td>106</td>
<td>$^{106}$Pd</td>
<td>105.903 478</td>
<td>27.33%</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Silver</td>
<td>107</td>
<td>$^{107}$Ag</td>
<td>106.905 092</td>
<td>51.84%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>109</td>
<td>$^{109}$Ag</td>
<td>108.904 757</td>
<td>48.16%</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Cadmium</td>
<td>114</td>
<td>$^{114}$Cd</td>
<td>113.903 357</td>
<td>28.73%</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Indium</td>
<td>115</td>
<td>$^{115}$In</td>
<td>114.903 880</td>
<td>95.7%, $\beta^-$</td>
<td>4.4×10$^{14}$ y</td>
</tr>
<tr>
<td>50</td>
<td>Tin</td>
<td>120</td>
<td>$^{120}$Sn</td>
<td>119.902 200</td>
<td>32.59%</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Antimony</td>
<td>121</td>
<td>$^{121}$Sb</td>
<td>120.903 821</td>
<td>57.3%</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Tellurium</td>
<td>130</td>
<td>$^{130}$Te</td>
<td>129.906 229</td>
<td>33.8%, $\beta^-$</td>
<td>2.5×10$^{21}$ y</td>
</tr>
<tr>
<td>53</td>
<td>Iodine</td>
<td>127</td>
<td>$^{127}$I</td>
<td>126.904 473</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>131</td>
<td>$^{131}$I</td>
<td>130.906 114</td>
<td>$\beta^-$</td>
<td>8.040 d</td>
</tr>
<tr>
<td>54</td>
<td>Xenon</td>
<td>132</td>
<td>$^{132}$Xe</td>
<td>131.904 144</td>
<td>26.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>136</td>
<td>$^{136}$Xe</td>
<td>135.907 214</td>
<td>8.9%</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Cesium</td>
<td>133</td>
<td>$^{133}$Cs</td>
<td>132.905 429</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>134</td>
<td>$^{134}$Cs</td>
<td>133.906 696</td>
<td>EC, $\beta^-$</td>
<td>2.06 y</td>
</tr>
<tr>
<td>56</td>
<td>Barium</td>
<td>137</td>
<td>$^{137}$Ba</td>
<td>136.905 812</td>
<td>11.23%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>138</td>
<td>$^{138}$Ba</td>
<td>137.905 232</td>
<td>71.70%</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>Lanthanum</td>
<td>139</td>
<td>$^{139}$La</td>
<td>138.906 346</td>
<td>99.91%</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>Cerium</td>
<td>140</td>
<td>$^{140}$Ce</td>
<td>139.905 433</td>
<td>88.48%</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>Praseodymium</td>
<td>141</td>
<td>$^{141}$Pr</td>
<td>140.907 647</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>Neodymium</td>
<td>142</td>
<td>$^{142}$Nd</td>
<td>141.907 719</td>
<td>27.13%</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>Promethium</td>
<td>145</td>
<td>$^{145}$Pm</td>
<td>144.912 743</td>
<td>EC, $\alpha$</td>
<td>17.7 y</td>
</tr>
<tr>
<td>62</td>
<td>Samarium</td>
<td>152</td>
<td>$^{152}$Sm</td>
<td>151.919 729</td>
<td>26.7%</td>
<td></td>
</tr>
<tr>
<td>Atomic Number, Z</td>
<td>Name</td>
<td>Atomic Mass Number, A</td>
<td>Symbol</td>
<td>Atomic Mass (u)</td>
<td>Percent Abundance or Decay Mode</td>
<td>Half-life, $t_{1/2}$</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>--------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>63</td>
<td>Europium</td>
<td>153</td>
<td>$^{153}\text{Eu}$</td>
<td>152.921 225</td>
<td>52.2%</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>Gadolinium</td>
<td>158</td>
<td>$^{158}\text{Gd}$</td>
<td>157.924 099</td>
<td>24.84%</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>Terbium</td>
<td>159</td>
<td>$^{159}\text{Tb}$</td>
<td>158.925 342</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>Dysprosium</td>
<td>164</td>
<td>$^{164}\text{Dy}$</td>
<td>163.929 171</td>
<td>28.2%</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>Holmium</td>
<td>165</td>
<td>$^{165}\text{Ho}$</td>
<td>164.930 319</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>Erbium</td>
<td>166</td>
<td>$^{166}\text{Er}$</td>
<td>165.930 290</td>
<td>33.6%</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>Thulium</td>
<td>169</td>
<td>$^{169}\text{Tm}$</td>
<td>168.934 212</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Ytterbium</td>
<td>174</td>
<td>$^{174}\text{Yb}$</td>
<td>173.938 859</td>
<td>31.8%</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Lutecium</td>
<td>175</td>
<td>$^{175}\text{Lu}$</td>
<td>174.940 770</td>
<td>97.41%</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>Hafnium</td>
<td>180</td>
<td>$^{180}\text{Hf}$</td>
<td>179.946 545</td>
<td>35.10%</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>Tantalum</td>
<td>181</td>
<td>$^{181}\text{Ta}$</td>
<td>180.947 992</td>
<td>99.98%</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>Tungsten</td>
<td>184</td>
<td>$^{184}\text{W}$</td>
<td>183.950 928</td>
<td>30.67%</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>Rhenium</td>
<td>187</td>
<td>$^{187}\text{Re}$</td>
<td>186.955 744</td>
<td>62.6%, $^\beta^-$</td>
<td>$4.6\times10^{10}$ y</td>
</tr>
<tr>
<td>76</td>
<td>Osmium</td>
<td>191</td>
<td>$^{191}\text{Os}$</td>
<td>190.960 920</td>
<td>$^\beta^-$</td>
<td>15.4 d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>192</td>
<td>$^{192}\text{Os}$</td>
<td>191.961 467</td>
<td>41.0%</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>Iridium</td>
<td>191</td>
<td>$^{191}\text{Ir}$</td>
<td>190.960 584</td>
<td>37.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>193</td>
<td>$^{193}\text{Ir}$</td>
<td>192.962 917</td>
<td>62.7%</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Platinum</td>
<td>195</td>
<td>$^{195}\text{Pt}$</td>
<td>194.964 766</td>
<td>33.8%</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Gold</td>
<td>197</td>
<td>$^{197}\text{Au}$</td>
<td>196.966 543</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>198</td>
<td>$^{198}\text{Au}$</td>
<td>197.968 217</td>
<td>$^\beta^-$</td>
<td>2.696 d</td>
</tr>
<tr>
<td>80</td>
<td>Mercury</td>
<td>199</td>
<td>$^{199}\text{Hg}$</td>
<td>198.968 253</td>
<td>16.87%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>202</td>
<td>$^{202}\text{Hg}$</td>
<td>201.970 617</td>
<td>29.86%</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>Thallium</td>
<td>205</td>
<td>$^{205}\text{Tl}$</td>
<td>204.974 401</td>
<td>70.48%</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Lead</td>
<td>206</td>
<td>$^{206}\text{Pb}$</td>
<td>205.974 440</td>
<td>24.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>207</td>
<td>$^{207}\text{Pb}$</td>
<td>206.975 872</td>
<td>22.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>208</td>
<td>$^{208}\text{Pb}$</td>
<td>207.976 627</td>
<td>52.4%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>210</td>
<td>$^{210}\text{Pb}$</td>
<td>209.984 163</td>
<td>$^\alpha$, $^\beta^-$</td>
<td>22.3 y</td>
</tr>
<tr>
<td></td>
<td></td>
<td>211</td>
<td>$^{211}\text{Pb}$</td>
<td>210.988 735</td>
<td>$^\beta^-$</td>
<td>36.1 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>212</td>
<td>$^{212}\text{Pb}$</td>
<td>211.991 871</td>
<td>$^\beta^-$</td>
<td>10.64 h</td>
</tr>
<tr>
<td>83</td>
<td>Bismuth</td>
<td>209</td>
<td>$^{209}\text{Bi}$</td>
<td>208.980 374</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>211</td>
<td>$^{211}\text{Bi}$</td>
<td>210.987 255</td>
<td>$^\alpha$, $^\beta^-$</td>
<td>2.14 min</td>
</tr>
<tr>
<td>84</td>
<td>Polonion</td>
<td>210</td>
<td>$^{210}\text{Po}$</td>
<td>209.982 848</td>
<td>$^\alpha$</td>
<td>138.38 d</td>
</tr>
<tr>
<td>85</td>
<td>Astatine</td>
<td>218</td>
<td>$^{218}\text{At}$</td>
<td>218.008 684</td>
<td>$^\alpha$, $^\beta^-$</td>
<td>1.6 s</td>
</tr>
<tr>
<td>Atomic Number, Z</td>
<td>Name</td>
<td>Atomic Mass Number, A</td>
<td>Symbol</td>
<td>Atomic Mass (u)</td>
<td>Percent Abundance or Decay Mode</td>
<td>Half-life, $t_{1/2}$</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>------------------------</td>
<td>--------</td>
<td>-----------------</td>
<td>---------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>86</td>
<td>Radon</td>
<td>222</td>
<td>$^{222}\text{Rn}$</td>
<td>222.017 570</td>
<td>$\alpha$</td>
<td>3.82 d</td>
</tr>
<tr>
<td>87</td>
<td>Francium</td>
<td>223</td>
<td>$^{223}\text{Fr}$</td>
<td>223.019 733</td>
<td>$\alpha, \beta^-$</td>
<td>21.8 min</td>
</tr>
<tr>
<td>88</td>
<td>Radium</td>
<td>226</td>
<td>$^{226}\text{Ra}$</td>
<td>226.025 402</td>
<td>$\alpha$</td>
<td>1.60 x 10^3 y</td>
</tr>
<tr>
<td>89</td>
<td>Actinium</td>
<td>227</td>
<td>$^{227}\text{Ac}$</td>
<td>227.027 750</td>
<td>$\alpha, \beta^-$</td>
<td>21.8 y</td>
</tr>
<tr>
<td>90</td>
<td>Thorium</td>
<td>228</td>
<td>$^{228}\text{Th}$</td>
<td>228.028 715</td>
<td>$\alpha$</td>
<td>1.91 y</td>
</tr>
<tr>
<td>232</td>
<td></td>
<td></td>
<td></td>
<td>232.038 054</td>
<td>$100%, \alpha$</td>
<td>1.41 x 10^10 y</td>
</tr>
<tr>
<td>91</td>
<td>Protactinium</td>
<td>231</td>
<td>$^{231}\text{Pa}$</td>
<td>231.035 880</td>
<td>$\alpha$</td>
<td>3.28 x 10^4 y</td>
</tr>
<tr>
<td>92</td>
<td>Uranium</td>
<td>233</td>
<td>$^{233}\text{U}$</td>
<td>233.039 628</td>
<td>$\alpha$</td>
<td>1.59 x 10^3 y</td>
</tr>
<tr>
<td>235</td>
<td></td>
<td></td>
<td></td>
<td>235.043 924</td>
<td>0.720%, $\alpha$</td>
<td>7.04 x 10^8 y</td>
</tr>
<tr>
<td>236</td>
<td></td>
<td></td>
<td></td>
<td>236.045 562</td>
<td>$\alpha$</td>
<td>2.34 x 10^7 y</td>
</tr>
<tr>
<td>238</td>
<td></td>
<td></td>
<td></td>
<td>238.050 784</td>
<td>99.2745%, $\alpha$</td>
<td>4.47 x 10^9 y</td>
</tr>
<tr>
<td>239</td>
<td></td>
<td></td>
<td></td>
<td>239.054 289</td>
<td>$\beta^-$</td>
<td>23.5 min</td>
</tr>
<tr>
<td>93</td>
<td>Neptunium</td>
<td>239</td>
<td>$^{239}\text{Np}$</td>
<td>239.052 933</td>
<td>$\beta^-$</td>
<td>2.355 d</td>
</tr>
<tr>
<td>94</td>
<td>Plutonium</td>
<td>239</td>
<td>$^{239}\text{Pu}$</td>
<td>239.052 157</td>
<td>$\alpha$</td>
<td>2.41 x 10^4 y</td>
</tr>
<tr>
<td>95</td>
<td>Americium</td>
<td>243</td>
<td>$^{243}\text{Am}$</td>
<td>243.061 375</td>
<td>$\alpha$, fission</td>
<td>7.37 x 10^3 y</td>
</tr>
<tr>
<td>96</td>
<td>Curium</td>
<td>245</td>
<td>$^{245}\text{Cm}$</td>
<td>245.065 483</td>
<td>$\alpha$</td>
<td>8.50 x 10^3 y</td>
</tr>
<tr>
<td>97</td>
<td>Berkelium</td>
<td>247</td>
<td>$^{247}\text{Bk}$</td>
<td>247.070 300</td>
<td>$\alpha$</td>
<td>1.38 x 10^3 y</td>
</tr>
<tr>
<td>98</td>
<td>Californium</td>
<td>249</td>
<td>$^{249}\text{Cf}$</td>
<td>249.074 844</td>
<td>$\alpha$</td>
<td>351 y</td>
</tr>
<tr>
<td>99</td>
<td>Einsteinium</td>
<td>254</td>
<td>$^{254}\text{Es}$</td>
<td>254.088 019</td>
<td>$\alpha, \beta^-$</td>
<td>276 d</td>
</tr>
<tr>
<td>100</td>
<td>Fermium</td>
<td>253</td>
<td>$^{253}\text{Fm}$</td>
<td>253.085 173</td>
<td>EC, $\alpha$</td>
<td>3.00 d</td>
</tr>
<tr>
<td>101</td>
<td>Mendelevium</td>
<td>255</td>
<td>$^{255}\text{Md}$</td>
<td>255.091 081</td>
<td>EC, $\alpha$</td>
<td>27 min</td>
</tr>
<tr>
<td>102</td>
<td>Nobelium</td>
<td>255</td>
<td>$^{255}\text{No}$</td>
<td>255.093 260</td>
<td>EC, $\alpha$</td>
<td>3.1 min</td>
</tr>
<tr>
<td>103</td>
<td>Lawrencium</td>
<td>257</td>
<td>$^{257}\text{Lr}$</td>
<td>257.099 480</td>
<td>EC, $\alpha$</td>
<td>0.646 s</td>
</tr>
<tr>
<td>104</td>
<td>Rutherfordium</td>
<td>261</td>
<td>$^{261}\text{Rf}$</td>
<td>261.108 690</td>
<td>$\alpha$</td>
<td>1.08 min</td>
</tr>
<tr>
<td>105</td>
<td>Dubnium</td>
<td>262</td>
<td>$^{262}\text{Db}$</td>
<td>262.113 760</td>
<td>$\alpha$, fission</td>
<td>34 s</td>
</tr>
<tr>
<td>106</td>
<td>Seaborgium</td>
<td>263</td>
<td>$^{263}\text{Sg}$</td>
<td>263.11 86</td>
<td>$\alpha$, fission</td>
<td>0.8 s</td>
</tr>
<tr>
<td>107</td>
<td>Bohrium</td>
<td>262</td>
<td>$^{262}\text{Bh}$</td>
<td>262.123 1</td>
<td>$\alpha$</td>
<td>0.102 s</td>
</tr>
<tr>
<td>108</td>
<td>Hassium</td>
<td>264</td>
<td>$^{264}\text{Hs}$</td>
<td>264.128 5</td>
<td>$\alpha$</td>
<td>0.08 ms</td>
</tr>
<tr>
<td>109</td>
<td>Meitnerium</td>
<td>266</td>
<td>$^{266}\text{Mt}$</td>
<td>266.137 8</td>
<td>$\alpha$</td>
<td>3.4 ms</td>
</tr>
</tbody>
</table>
Decay modes are $\alpha$, $\beta^-$, $\beta^+$, electron capture (EC) and isomeric transition (IT). EC results in the same daughter nucleus as would $\beta^+$ decay. IT is a transition from a metastable excited state. Energies for $\beta^\pm$ decays are the maxima; average energies are roughly one-half the maxima.
<table>
<thead>
<tr>
<th>Isotope</th>
<th>$t_{1/2}$</th>
<th>DecayMode(s)</th>
<th>Energy(MeV)</th>
<th>Percent</th>
<th>$\gamma$-Ray Energy(MeV)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^3$H</td>
<td>12.33 y</td>
<td>$\beta^-$</td>
<td>0.0186</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{14}$C</td>
<td>5730 y</td>
<td>$\beta^-$</td>
<td>0.156</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{13}$N</td>
<td>9.96 min</td>
<td>$\beta^+$</td>
<td>1.20</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{22}$Na</td>
<td>2.602 y</td>
<td>$\beta^+$</td>
<td>0.55</td>
<td>90%</td>
<td>$\gamma$</td>
<td>1.27</td>
</tr>
<tr>
<td>$^{32}$P</td>
<td>14.28 d</td>
<td>$\beta^-$</td>
<td>1.71</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{35}$S</td>
<td>87.4 d</td>
<td>$\beta^-$</td>
<td>0.167</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{36}$Cl</td>
<td>3.00x10$^5$ y</td>
<td>$\beta^-$</td>
<td>0.710</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{40}$K</td>
<td>1.28x10$^9$ y</td>
<td>$\beta^-$</td>
<td>1.31</td>
<td>89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{43}$K</td>
<td>22.3 h</td>
<td>$\beta^-$</td>
<td>0.827</td>
<td>87%</td>
<td>$\gamma$</td>
<td>0.373</td>
</tr>
<tr>
<td>$^{45}$Ca</td>
<td>165 d</td>
<td>$\beta^-$</td>
<td>0.257</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{51}$Cr</td>
<td>27.70 d</td>
<td>EC</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>0.320</td>
</tr>
<tr>
<td>$^{52}$Mn</td>
<td>5.59d</td>
<td>$\beta^+$</td>
<td>3.69</td>
<td>28%</td>
<td>$\gamma$</td>
<td>1.33</td>
</tr>
<tr>
<td>$^{52}$Fe</td>
<td>8.27 h</td>
<td>$\beta^+$</td>
<td>1.80</td>
<td>43%</td>
<td></td>
<td>0.169</td>
</tr>
<tr>
<td>$^{59}$Fe</td>
<td>44.6 d</td>
<td>$\beta^-$ s</td>
<td>0.273</td>
<td>45%</td>
<td>$\gamma$</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.466</td>
<td>55%</td>
<td></td>
<td>1.29</td>
</tr>
<tr>
<td>$^{60}$Co</td>
<td>5.271 y</td>
<td>$\beta^-$</td>
<td>0.318</td>
<td>100%</td>
<td>$\gamma$</td>
<td>1.17</td>
</tr>
<tr>
<td>$^{65}$Zn</td>
<td>244.1 d</td>
<td>EC</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>1.12</td>
</tr>
<tr>
<td>$^{67}$Ga</td>
<td>78.3 h</td>
<td>EC</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>0.0933</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>others</td>
<td></td>
</tr>
<tr>
<td>$^{75}$Se</td>
<td>118.5 d</td>
<td>EC</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>others</td>
<td></td>
</tr>
<tr>
<td>$^{80}$Rb</td>
<td>18.8 d</td>
<td>$\beta^-$ s</td>
<td>0.69</td>
<td>9%</td>
<td>$\gamma$</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.77</td>
<td>91%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{85}$Sr</td>
<td>64.8 d</td>
<td>EC</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>0.514</td>
</tr>
<tr>
<td>$^{90}$Sr</td>
<td>28.8 y</td>
<td>$\beta^-$</td>
<td>0.546</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{90}$Y</td>
<td>64.1 h</td>
<td>$\beta^-$</td>
<td>2.28</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{99m}$Tc</td>
<td>6.02 h</td>
<td>IT</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>0.142</td>
</tr>
<tr>
<td>Isotope</td>
<td>$t_{1/2}$</td>
<td>DecayMode(s)</td>
<td>Energy(MeV)</td>
<td>Percent</td>
<td>$\gamma$-Ray Energy(MeV)</td>
<td>Percent</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------</td>
<td>--------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>$^{113m}$In</td>
<td>99.5 min</td>
<td>IT</td>
<td></td>
<td></td>
<td>$\gamma$</td>
<td>0.392</td>
</tr>
<tr>
<td>$^{123}$I</td>
<td>13.0 h</td>
<td>EC</td>
<td>0.248</td>
<td>7%</td>
<td>$\gamma^s$</td>
<td>0.364</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.607</td>
<td>93%</td>
<td>others</td>
<td></td>
</tr>
<tr>
<td>$^{129}$Cs</td>
<td>32.3 h</td>
<td>EC</td>
<td></td>
<td></td>
<td>$\gamma^s$</td>
<td>0.0400</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.372</td>
<td>32%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.411</td>
<td>25%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>others</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{131}$I</td>
<td>8.040 d</td>
<td>$\beta^-$ s</td>
<td>0.511</td>
<td>95%</td>
<td>$\gamma$</td>
<td>0.662</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.17</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{137}$Cs</td>
<td>30.17 y</td>
<td>$\beta^-$ s</td>
<td>1.035</td>
<td>$\approx$100%</td>
<td>$\gamma^s$</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.44</td>
<td>65%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.537</td>
<td>24%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>others</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{140}$Ba</td>
<td>12.79 d</td>
<td>$\beta^-$</td>
<td>1.161</td>
<td>$\approx$100%</td>
<td>$\gamma$</td>
<td>0.0733</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.87</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{198}$Au</td>
<td>2.696 d</td>
<td>$\beta^-$</td>
<td>4.68</td>
<td>5%</td>
<td>$\gamma$</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.87</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{197}$Hg</td>
<td>64.1 h</td>
<td>EC</td>
<td>5.41</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.87</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{210}$Po</td>
<td>138.38 d</td>
<td>$\alpha$</td>
<td>4.68</td>
<td>$\approx$100%</td>
<td>$\gamma^s$</td>
<td>numerous</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.27</td>
<td>77%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{226}$Ra</td>
<td>$1.60 \times 10^3$ y</td>
<td>$\alpha s$</td>
<td>4.22</td>
<td>23%</td>
<td>$\gamma$</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.27</td>
<td>77%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{235}$U</td>
<td>$7.038 \times 10^8$ y</td>
<td>$\alpha$</td>
<td>4.68</td>
<td>$\approx$100%</td>
<td>$\gamma^s$</td>
<td>numerous</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.96 (max.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{238}$U</td>
<td>$4.468 \times 10^9$ y</td>
<td>$\alpha s$</td>
<td>4.22</td>
<td>23%</td>
<td>$\gamma$</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.27</td>
<td>77%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{237}$Np</td>
<td>$2.14 \times 10^6$ y</td>
<td>$\alpha s$</td>
<td>numerous</td>
<td></td>
<td>$\gamma^s$</td>
<td>numerous</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.96 (max.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{239}$Pu</td>
<td>$2.41 \times 10^4$ y</td>
<td>$\alpha s$</td>
<td>5.19</td>
<td>11%</td>
<td>$\gamma^s$</td>
<td>$7.5 \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.23</td>
<td>15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.24</td>
<td>73%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>others</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{243}$Am</td>
<td>$7.37 \times 10^3$ y</td>
<td>$\alpha s$</td>
<td>Max. 5.44</td>
<td></td>
<td>$\gamma^s$</td>
<td>0.075</td>
</tr>
</tbody>
</table>
This appendix is broken into several tables.

- **Table C1,** Important Constants
- **Table C2,** Submicroscopic Masses
- **Table C3,** Solar System Data
- **Table C4,** Metric Prefixes for Powers of Ten and Their Symbols
- **Table C5,** The Greek Alphabet
- **Table C6,** SI units
- **Table C7,** Selected British Units
- **Table C8,** Other Units
- **Table C9,** Useful Formulae

### Table C1 Important Constants

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Best Value</th>
<th>Approximate Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c )</td>
<td>Speed of light in vacuum</td>
<td>( 2.99792458 \times 10^8 \text{ m/s} )</td>
<td>( 3.00 \times 10^8 \text{ m/s} )</td>
</tr>
<tr>
<td>( G )</td>
<td>Gravitational constant</td>
<td>( 6.67384(80) \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2 )</td>
<td>( 6.67 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2 )</td>
</tr>
<tr>
<td>( N_A )</td>
<td>Avogadro's number</td>
<td>( 6.02214129(27) \times 10^{23} )</td>
<td>( 6.02 \times 10^{23} )</td>
</tr>
<tr>
<td>( k )</td>
<td>Boltzmann's constant</td>
<td>( 1.3806488(13) \times 10^{-23} \text{ J/K} )</td>
<td>( 1.38 \times 10^{-23} \text{ J/K} )</td>
</tr>
<tr>
<td>( R )</td>
<td>Gas constant</td>
<td>( 8.3144621(75) \text{ J/mol} \cdot \text{K} )</td>
<td>( 8.31 \text{ J/mol} \cdot \text{K} = 1.99 \text{ cal/mol} \cdot \text{K} = 0.0821 \text{ atm} \cdot \text{L} \cdot \text{mol} \cdot \text{K} )</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>Stefan-Boltzmann constant</td>
<td>( 5.670373(21) \times 10^{-8} \text{ W} \cdot \text{m}^2/\text{K} )</td>
<td>( 5.67 \times 10^{-8} \text{ W} \cdot \text{m}^2/\text{K} )</td>
</tr>
<tr>
<td>( k )</td>
<td>Coulomb force constant</td>
<td>( 8.987551788 \ldots \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2 )</td>
<td>( 8.99 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2 )</td>
</tr>
<tr>
<td>( q_e )</td>
<td>Charge on electron</td>
<td>( -1.602176565(35) \times 10^{-19} \text{ C} )</td>
<td>( -1.60 \times 10^{-19} \text{ C} )</td>
</tr>
<tr>
<td>( \varepsilon_0 )</td>
<td>Permittivity of free space</td>
<td>( 8.854187817 \ldots \times 10^{-12} \text{ C}^2/\text{N} \cdot \text{m}^2 )</td>
<td>( 8.85 \times 10^{-12} \text{ C}^2/\text{N} \cdot \text{m}^2 )</td>
</tr>
<tr>
<td>( \mu_0 )</td>
<td>Permeability of free space</td>
<td>( 4\pi \times 10^{-7} \text{ T} \cdot \text{m}/\text{A} )</td>
<td>( 1.26 \times 10^{-6} \text{ T} \cdot \text{m}/\text{A} )</td>
</tr>
<tr>
<td>( h )</td>
<td>Planck's constant</td>
<td>( 6.62606957(29) \times 10^{-34} \text{ J} \cdot \text{s} )</td>
<td>( 6.63 \times 10^{-34} \text{ J} \cdot \text{s} )</td>
</tr>
</tbody>
</table>

### Table C2 Submicroscopic Masses

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Best Value</th>
<th>Approximate Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_e )</td>
<td>Electron mass</td>
<td>( 9.10938291(40) \times 10^{-31} \text{ kg} )</td>
<td>( 9.11 \times 10^{-31} \text{ kg} )</td>
</tr>
<tr>
<td>( m_p )</td>
<td>Proton mass</td>
<td>( 1.672621777(74) \times 10^{-27} \text{ kg} )</td>
<td>( 1.6726 \times 10^{-27} \text{ kg} )</td>
</tr>
<tr>
<td>( m_n )</td>
<td>Neutron mass</td>
<td>( 1.674927351(74) \times 10^{-27} \text{ kg} )</td>
<td>( 1.6749 \times 10^{-27} \text{ kg} )</td>
</tr>
<tr>
<td>( u )</td>
<td>Atomic mass unit</td>
<td>( 1.660538921(73) \times 10^{-27} \text{ kg} )</td>
<td>( 1.6605 \times 10^{-27} \text{ kg} )</td>
</tr>
</tbody>
</table>

---

1. Stated values are according to the National Institute of Standards and Technology Reference on Constants, Units, and Uncertainty, [www.physics.nist.gov/cuu](http://www.physics.nist.gov/cuu) (accessed May 18, 2012). Values in parentheses are the uncertainties in the last digits. Numbers without uncertainties are exact as defined.
2. Stated values are according to the National Institute of Standards and Technology Reference on Constants, Units, and Uncertainty, [www.physics.nist.gov/cuu](http://www.physics.nist.gov/cuu) (accessed May 18, 2012). Values in parentheses are the uncertainties in the last digits. Numbers without uncertainties are exact as defined.
### Table C3 Solar System Data

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun</td>
<td>mass</td>
<td>$1.99 \times 10^{30} \text{kg}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>average radius</td>
<td>$6.96 \times 10^8 \text{m}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Earth</td>
<td>mass</td>
<td>$5.9736 \times 10^{24} \text{kg}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>average radius</td>
<td>$6.376 \times 10^6 \text{m}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>orbital period</td>
<td>$3.16 \times 10^7 \text{s}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moon</td>
<td>mass</td>
<td>$7.35 \times 10^{22} \text{kg}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>average radius</td>
<td>$1.74 \times 10^6 \text{m}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>orbital period (average)</td>
<td>$2.36 \times 10^6 \text{s}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Earth-moon distance (average)</td>
<td>$3.84 \times 10^8 \text{m}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table C4 Metric Prefixes for Powers of Ten and Their Symbols

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Symbol</th>
<th>Value</th>
<th>Prefix</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>tera</td>
<td>T</td>
<td>$10^{12}$</td>
<td>deci</td>
<td>d</td>
<td>$10^{-1}$</td>
</tr>
<tr>
<td>giga</td>
<td>G</td>
<td>$10^{9}$</td>
<td>centi</td>
<td>c</td>
<td>$10^{-2}$</td>
</tr>
<tr>
<td>mega</td>
<td>M</td>
<td>$10^{6}$</td>
<td>milli</td>
<td>m</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>kilo</td>
<td>k</td>
<td>$10^{3}$</td>
<td>micro</td>
<td>µ</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>hecto</td>
<td>h</td>
<td>$10^{2}$</td>
<td>nano</td>
<td>n</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>deka</td>
<td>da</td>
<td>$10^{1}$</td>
<td>pico</td>
<td>p</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>$10^{0}$</td>
<td>( = 1)</td>
<td>femto</td>
<td>$10^{-15}$</td>
</tr>
</tbody>
</table>

### Table C5 The Greek Alphabet

<table>
<thead>
<tr>
<th>Alpha</th>
<th>Α</th>
<th>α</th>
<th>Eta</th>
<th>Η</th>
<th>η</th>
<th>Nu</th>
<th>Ν</th>
<th>ν</th>
<th>Tau</th>
<th>Τ</th>
<th>τ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>Β</td>
<td>β</td>
<td>Theta</td>
<td>Θ</td>
<td>θ</td>
<td>Xi</td>
<td>Ξ</td>
<td>ξ</td>
<td>Upsilon</td>
<td>Υ</td>
<td>υ</td>
</tr>
<tr>
<td>Gamma</td>
<td>Γ</td>
<td>γ</td>
<td>Iota</td>
<td>Ι</td>
<td>ι</td>
<td>Omicron</td>
<td>Ο</td>
<td>ο</td>
<td>Phi</td>
<td>Φ</td>
<td>ϕ</td>
</tr>
<tr>
<td>Delta</td>
<td>Δ</td>
<td>δ</td>
<td>Kappa</td>
<td>Κ</td>
<td>κ</td>
<td>Pi</td>
<td>Π</td>
<td>π</td>
<td>Chi</td>
<td>Χ</td>
<td>χ</td>
</tr>
<tr>
<td>Epsilon</td>
<td>Ε</td>
<td>ε</td>
<td>Lambda</td>
<td>Λ</td>
<td>λ</td>
<td>Rho</td>
<td>Ρ</td>
<td>ρ</td>
<td>Psi</td>
<td>Ψ</td>
<td>ψ</td>
</tr>
<tr>
<td>Zeta</td>
<td>Ζ</td>
<td>ζ</td>
<td>Mu</td>
<td>Μ</td>
<td>μ</td>
<td>Sigma</td>
<td>Σ</td>
<td>σ</td>
<td>Omega</td>
<td>Ω</td>
<td>ω</td>
</tr>
</tbody>
</table>
### Table C6 SI Units

<table>
<thead>
<tr>
<th>Entity</th>
<th>Abbreviation</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fundamental units</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>m</td>
<td>meter</td>
</tr>
<tr>
<td>Mass</td>
<td>kg</td>
<td>kilogram</td>
</tr>
<tr>
<td>Time</td>
<td>s</td>
<td>second</td>
</tr>
<tr>
<td>Current</td>
<td>A</td>
<td>ampere</td>
</tr>
<tr>
<td><strong>Supplementary unit</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angle</td>
<td>rad</td>
<td>radian</td>
</tr>
<tr>
<td><strong>Derived units</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Force</td>
<td>N</td>
<td>newton</td>
</tr>
<tr>
<td>Energy</td>
<td>J</td>
<td>joule</td>
</tr>
<tr>
<td>Power</td>
<td>W</td>
<td>watt</td>
</tr>
<tr>
<td>Pressure</td>
<td>Pa</td>
<td>pascal</td>
</tr>
<tr>
<td>Frequency</td>
<td>Hz</td>
<td>hertz</td>
</tr>
<tr>
<td>Electronic potential</td>
<td>V = J/C</td>
<td>volt</td>
</tr>
<tr>
<td>Capacitance</td>
<td>F = C/V</td>
<td>farad</td>
</tr>
<tr>
<td>Charge</td>
<td>C = s ⋅ A</td>
<td>coulomb</td>
</tr>
<tr>
<td>Resistance</td>
<td>Ω = V/A</td>
<td>ohm</td>
</tr>
<tr>
<td>Magnetic field</td>
<td>T = N/(A ⋅ m)</td>
<td>tesla</td>
</tr>
<tr>
<td>Nuclear decay rate</td>
<td>Bq = 1/s</td>
<td>becquerel</td>
</tr>
</tbody>
</table>

### Table C7 Selected British Units

<table>
<thead>
<tr>
<th>Entity</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>1 inch (in.) = 2.54 cm (exactly)</td>
</tr>
<tr>
<td></td>
<td>1 foot (ft) = 0.3048 m</td>
</tr>
<tr>
<td></td>
<td>1 mile (mi) = 1.609 km</td>
</tr>
<tr>
<td>Force</td>
<td>1 pound (lb) = 4.448 N</td>
</tr>
<tr>
<td>Energy</td>
<td>1 British thermal unit (Btu) = 1.055×10³ J</td>
</tr>
<tr>
<td>Power</td>
<td>1 horsepower (hp) = 746 W</td>
</tr>
<tr>
<td>Pressure</td>
<td>1 lb/in² = 6.895×10³ Pa</td>
</tr>
<tr>
<td>Table C8 Other Units</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Length</strong></td>
<td>1 light year (ly) = 9.46×10^{15} m</td>
</tr>
<tr>
<td>1 astronomical unit (au) = 1.50×10^{11} m</td>
<td></td>
</tr>
<tr>
<td>1 nautical mile = 1.852 km</td>
<td></td>
</tr>
<tr>
<td>1 angstrom(Å) = 10^{-10} m</td>
<td></td>
</tr>
<tr>
<td><strong>Area</strong></td>
<td>1 acre (ac) = 4.05×10^{3} m^2</td>
</tr>
<tr>
<td>1 square foot (ft^2) = 9.29×10^{-2} m^2</td>
<td></td>
</tr>
<tr>
<td>1 barn (b) = 10^{-28} m^2</td>
<td></td>
</tr>
<tr>
<td><strong>Volume</strong></td>
<td>1 liter (L) = 10^{-3} m^3</td>
</tr>
<tr>
<td>1 U.S. gallon (gal) = 3.785×10^{-3} m^3</td>
<td></td>
</tr>
<tr>
<td><strong>Mass</strong></td>
<td>1 solar mass = 1.99×10^{30} kg</td>
</tr>
<tr>
<td>1 metric ton = 10^3 kg</td>
<td></td>
</tr>
<tr>
<td>1 atomic mass unit (u) = 1.6605×10^{-27} kg</td>
<td></td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td>1 year (y) = 3.16×10^7 s</td>
</tr>
<tr>
<td>1 day (d) = 86,400 s</td>
<td></td>
</tr>
<tr>
<td><strong>Speed</strong></td>
<td>1 mile per hour (mph) = 1.609 km/h</td>
</tr>
<tr>
<td>1 nautical mile per hour (naut) = 1.852 km/h</td>
<td></td>
</tr>
<tr>
<td><strong>Angle</strong></td>
<td>1 degree (°) = 1.745×10^{-2} rad</td>
</tr>
<tr>
<td>1 minute of arc (') = 1 / 60 degree</td>
<td></td>
</tr>
<tr>
<td>1 second of arc (&quot;) = 1 / 60 minute of arc</td>
<td></td>
</tr>
<tr>
<td>1 grad = 1.571×10^{-2} rad</td>
<td></td>
</tr>
<tr>
<td><strong>Energy</strong></td>
<td>1 kiloton TNT (kT) = 4.2×10^{12} J</td>
</tr>
<tr>
<td>1 kilowatt hour (kW·h) = 3.60×10^6 J</td>
<td></td>
</tr>
<tr>
<td>1 food calorie (kcal) = 4186 J</td>
<td></td>
</tr>
<tr>
<td>1 calorie (cal) = 4.186 J</td>
<td></td>
</tr>
<tr>
<td>1 electron volt (eV) = 1.60×10^{-19} J</td>
<td></td>
</tr>
<tr>
<td><strong>Pressure</strong></td>
<td>1 atmosphere (atm) = 1.013×10^5 Pa</td>
</tr>
<tr>
<td>1 millimeter of mercury (mm Hg) = 133.3 Pa</td>
<td></td>
</tr>
<tr>
<td>1 torricelli (torr) = 1 mm Hg = 133.3 Pa</td>
<td></td>
</tr>
<tr>
<td><strong>Nuclear decay rate</strong></td>
<td>1 curie (Ci) = 3.70×10^{10} Bq</td>
</tr>
</tbody>
</table>
### Table C9 Useful Formulae

<table>
<thead>
<tr>
<th>Formula</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumference of a circle with radius ( r ) or diameter ( d )</td>
<td>( C = 2\pi r = \pi d )</td>
</tr>
<tr>
<td>Area of a circle with radius ( r ) or diameter ( d )</td>
<td>( A = \pi r^2 = \pi d^2 / 4 )</td>
</tr>
<tr>
<td>Area of a sphere with radius ( r )</td>
<td>( A = 4\pi r^2 )</td>
</tr>
<tr>
<td>Volume of a sphere with radius ( r )</td>
<td>( V = (4 / 3)(\pi r^3) )</td>
</tr>
</tbody>
</table>

3. Stated values are according to the National Institute of Standards and Technology Reference on Constants, Units, and Uncertainty, [www.physics.nist.gov/cuu](http://www.physics.nist.gov/cuu) (accessed May 18, 2012). Values in parentheses are the uncertainties in the last digits. Numbers without uncertainties are exact as defined.

4. Stated values are according to the National Institute of Standards and Technology Reference on Constants, Units, and Uncertainty, [www.physics.nist.gov/cuu](http://www.physics.nist.gov/cuu) (accessed May 18, 2012). Values in parentheses are the uncertainties in the last digits. Numbers without uncertainties are exact as defined.
In this glossary, key symbols and notation are briefly defined.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>any symbol</td>
<td>average (indicated by a bar over a symbol—e.g., $\bar{v}$ is average velocity)</td>
</tr>
<tr>
<td>°C</td>
<td>Celsius degree</td>
</tr>
<tr>
<td>°F</td>
<td>Fahrenheit degree</td>
</tr>
<tr>
<td>//</td>
<td>parallel</td>
</tr>
<tr>
<td>⊥</td>
<td>perpendicular</td>
</tr>
<tr>
<td>α</td>
<td>proportional to</td>
</tr>
<tr>
<td>±</td>
<td>plus or minus</td>
</tr>
<tr>
<td>0</td>
<td>zero as a subscript denotes an initial value</td>
</tr>
<tr>
<td>α</td>
<td>alpha rays</td>
</tr>
<tr>
<td>α</td>
<td>angular acceleration</td>
</tr>
<tr>
<td>α</td>
<td>temperature coefficient(s) of resistivity</td>
</tr>
<tr>
<td>β</td>
<td>beta rays</td>
</tr>
<tr>
<td>β</td>
<td>sound level</td>
</tr>
<tr>
<td>β</td>
<td>volume coefficient of expansion</td>
</tr>
<tr>
<td>β−</td>
<td>electron emitted in nuclear beta decay</td>
</tr>
<tr>
<td>β+</td>
<td>positron decay</td>
</tr>
<tr>
<td>γ</td>
<td>gamma rays</td>
</tr>
<tr>
<td>γ</td>
<td>surface tension</td>
</tr>
<tr>
<td>$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$</td>
<td>a constant used in relativity</td>
</tr>
<tr>
<td>Δ</td>
<td>change in whatever quantity follows</td>
</tr>
<tr>
<td>δ</td>
<td>uncertainty in whatever quantity follows</td>
</tr>
<tr>
<td>ΔE</td>
<td>change in energy between the initial and final orbits of an electron in an atom</td>
</tr>
<tr>
<td>ΔE</td>
<td>uncertainty in energy</td>
</tr>
<tr>
<td>Δm</td>
<td>difference in mass between initial and final products</td>
</tr>
<tr>
<td>ΔN</td>
<td>number of decays that occur</td>
</tr>
<tr>
<td>Δp</td>
<td>change in momentum</td>
</tr>
<tr>
<td>Δp</td>
<td>uncertainty in momentum</td>
</tr>
<tr>
<td>ΔPEg</td>
<td>change in gravitational potential energy</td>
</tr>
<tr>
<td>Δθ</td>
<td>rotation angle</td>
</tr>
<tr>
<td>Δs</td>
<td>distance traveled along a circular path</td>
</tr>
<tr>
<td>Δt</td>
<td>uncertainty in time</td>
</tr>
<tr>
<td>Δt₀</td>
<td>proper time as measured by an observer at rest relative to the process</td>
</tr>
<tr>
<td>ΔV</td>
<td>potential difference</td>
</tr>
<tr>
<td>Δx</td>
<td>uncertainty in position</td>
</tr>
<tr>
<td>$\varepsilon_0$</td>
<td>permittivity of free space</td>
</tr>
<tr>
<td>η</td>
<td>viscosity</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$\theta$</td>
<td>angle between the force vector and the displacement vector</td>
</tr>
<tr>
<td>$\theta$</td>
<td>angle between two lines</td>
</tr>
<tr>
<td>$\theta$</td>
<td>contact angle</td>
</tr>
<tr>
<td>$\theta$</td>
<td>direction of the resultant</td>
</tr>
<tr>
<td>$\theta_b$</td>
<td>Brewster's angle</td>
</tr>
<tr>
<td>$\theta_c$</td>
<td>critical angle</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>dielectric constant</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>decay constant of a nuclide</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>wavelength</td>
</tr>
<tr>
<td>$\lambda_n$</td>
<td>wavelength in a medium</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>permeability of free space</td>
</tr>
<tr>
<td>$\mu_k$</td>
<td>coefficient of kinetic friction</td>
</tr>
<tr>
<td>$\mu_s$</td>
<td>coefficient of static friction</td>
</tr>
<tr>
<td>$\nu_e$</td>
<td>electron neutrino</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>positive pion</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>negative pion</td>
</tr>
<tr>
<td>$\pi^0$</td>
<td>neutral pion</td>
</tr>
<tr>
<td>$\rho$</td>
<td>density</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>critical density, the density needed to just halt universal expansion</td>
</tr>
<tr>
<td>$\rho_{fl}$</td>
<td>fluid density</td>
</tr>
<tr>
<td>$\rho_{obj}$</td>
<td>average density of an object</td>
</tr>
<tr>
<td>$\rho / \rho_w$</td>
<td>specific gravity</td>
</tr>
<tr>
<td>$\tau$</td>
<td>characteristic time constant for a resistance and inductance ($RL$) or resistance and capacitance ($RC$) circuit</td>
</tr>
<tr>
<td>$\tau$</td>
<td>characteristic time for a resistor and capacitor ($RC$) circuit</td>
</tr>
<tr>
<td>$\tau$</td>
<td>torque</td>
</tr>
<tr>
<td>$\Upsilon$</td>
<td>upsilon meson</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>magnetic flux</td>
</tr>
<tr>
<td>$\phi$</td>
<td>phase angle</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>ohm (unit)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>angular velocity</td>
</tr>
<tr>
<td>$A$</td>
<td>ampere (current unit)</td>
</tr>
<tr>
<td>$A$</td>
<td>area</td>
</tr>
<tr>
<td>$A$</td>
<td>cross-sectional area</td>
</tr>
<tr>
<td>$A$</td>
<td>total number of nucleons</td>
</tr>
<tr>
<td>$a$</td>
<td>acceleration</td>
</tr>
<tr>
<td>$a_B$</td>
<td>Bohr radius</td>
</tr>
<tr>
<td>$a_c$</td>
<td>centripetal acceleration</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>$\alpha_t$</td>
<td>tangential acceleration</td>
</tr>
<tr>
<td>AC</td>
<td>alternating current</td>
</tr>
<tr>
<td>AM</td>
<td>amplitude modulation</td>
</tr>
<tr>
<td>atm</td>
<td>atmosphere</td>
</tr>
<tr>
<td>$B$</td>
<td>baryon number</td>
</tr>
<tr>
<td>$B$</td>
<td>blue quark color</td>
</tr>
<tr>
<td>$\bar{B}$</td>
<td>antiblue (yellow) antiquark color</td>
</tr>
<tr>
<td>$b$</td>
<td>quark flavor bottom or beauty</td>
</tr>
<tr>
<td>$B$</td>
<td>bulk modulus</td>
</tr>
<tr>
<td>$B$</td>
<td>magnetic field strength</td>
</tr>
<tr>
<td>$B_{\text{int}}$</td>
<td>electron’s intrinsic magnetic field</td>
</tr>
<tr>
<td>$B_{\text{orb}}$</td>
<td>orbital magnetic field</td>
</tr>
<tr>
<td>BE</td>
<td>binding energy of a nucleus—it is the energy required to completely disassemble it into separate protons and neutrons</td>
</tr>
<tr>
<td>BE / $A$</td>
<td>binding energy per nucleon</td>
</tr>
<tr>
<td>Bq</td>
<td>becquerel—one decay per second</td>
</tr>
<tr>
<td>C</td>
<td>capacitance (amount of charge stored per volt)</td>
</tr>
<tr>
<td>C</td>
<td>coulomb (a fundamental SI unit of charge)</td>
</tr>
<tr>
<td>$C_p$</td>
<td>total capacitance in parallel</td>
</tr>
<tr>
<td>$C_s$</td>
<td>total capacitance in series</td>
</tr>
<tr>
<td>CG</td>
<td>center of gravity</td>
</tr>
<tr>
<td>CM</td>
<td>center of mass</td>
</tr>
<tr>
<td>c</td>
<td>quark flavor charm</td>
</tr>
<tr>
<td>c</td>
<td>specific heat</td>
</tr>
<tr>
<td>c</td>
<td>speed of light</td>
</tr>
<tr>
<td>Cal</td>
<td>kilocalorie</td>
</tr>
<tr>
<td>cal</td>
<td>calorie</td>
</tr>
<tr>
<td>COP$_{\text{hp}}$</td>
<td>heat pump’s coefficient of performance</td>
</tr>
<tr>
<td>COP$_{\text{ref}}$</td>
<td>coefficient of performance for refrigerators and air conditioners</td>
</tr>
<tr>
<td>$\cos \theta$</td>
<td>cosine</td>
</tr>
<tr>
<td>$\cot \theta$</td>
<td>cotangent</td>
</tr>
<tr>
<td>$\csc \theta$</td>
<td>cosecant</td>
</tr>
<tr>
<td>D</td>
<td>diffusion constant</td>
</tr>
<tr>
<td>d</td>
<td>displacement</td>
</tr>
<tr>
<td>d</td>
<td>quark flavor down</td>
</tr>
<tr>
<td>dB</td>
<td>decibel</td>
</tr>
<tr>
<td>$d_i$</td>
<td>distance of an image from the center of a lens</td>
</tr>
<tr>
<td>$d_o$</td>
<td>distance of an object from the center of a lens</td>
</tr>
<tr>
<td>DC</td>
<td>direct current</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>E</td>
<td>electric field strength</td>
</tr>
<tr>
<td>ε</td>
<td>emf (voltage) or Hall electromotive force</td>
</tr>
<tr>
<td>emf</td>
<td>electromotive force</td>
</tr>
<tr>
<td>E</td>
<td>energy of a single photon</td>
</tr>
<tr>
<td>E</td>
<td>nuclear reaction energy</td>
</tr>
<tr>
<td>E</td>
<td>relativistic total energy</td>
</tr>
<tr>
<td>E</td>
<td>total energy</td>
</tr>
<tr>
<td>E₀</td>
<td>ground state energy for hydrogen</td>
</tr>
<tr>
<td>E₀</td>
<td>rest energy</td>
</tr>
<tr>
<td>EC</td>
<td>electron capture</td>
</tr>
<tr>
<td>E_{cap}</td>
<td>energy stored in a capacitor</td>
</tr>
<tr>
<td>Eff</td>
<td>efficiency—the useful work output divided by the energy input</td>
</tr>
<tr>
<td>Eff_{C}</td>
<td>Carnot efficiency</td>
</tr>
<tr>
<td>E_{in}</td>
<td>energy consumed (food digested in humans)</td>
</tr>
<tr>
<td>E_{ind}</td>
<td>energy stored in an inductor</td>
</tr>
<tr>
<td>E_{out}</td>
<td>energy output</td>
</tr>
<tr>
<td>e</td>
<td>emissivity of an object</td>
</tr>
<tr>
<td>e⁺</td>
<td>antielectron or positron</td>
</tr>
<tr>
<td>eV</td>
<td>electron volt</td>
</tr>
<tr>
<td>F</td>
<td>farad (unit of capacitance, a coulomb per volt)</td>
</tr>
<tr>
<td>F</td>
<td>focal point of a lens</td>
</tr>
<tr>
<td>F</td>
<td>force</td>
</tr>
<tr>
<td>F</td>
<td>magnitude of a force</td>
</tr>
<tr>
<td>F</td>
<td>restoring force</td>
</tr>
<tr>
<td>F_B</td>
<td>buoyant force</td>
</tr>
<tr>
<td>F_c</td>
<td>centripetal force</td>
</tr>
<tr>
<td>F_i</td>
<td>force input</td>
</tr>
<tr>
<td>F_{net}</td>
<td>net force</td>
</tr>
<tr>
<td>F₀</td>
<td>force output</td>
</tr>
<tr>
<td>FM</td>
<td>frequency modulation</td>
</tr>
<tr>
<td>f</td>
<td>focal length</td>
</tr>
<tr>
<td>f</td>
<td>frequency</td>
</tr>
<tr>
<td>f₀</td>
<td>resonant frequency of a resistance, inductance, and capacitance (RLC) series circuit</td>
</tr>
<tr>
<td>f₀</td>
<td>threshold frequency for a particular material (photoelectric effect)</td>
</tr>
<tr>
<td>f₁</td>
<td>fundamental</td>
</tr>
<tr>
<td>f₂</td>
<td>first overtone</td>
</tr>
<tr>
<td>f₃</td>
<td>second overtone</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$f_B$</td>
<td>beat frequency</td>
</tr>
<tr>
<td>$f_k$</td>
<td>magnitude of kinetic friction</td>
</tr>
<tr>
<td>$f_s$</td>
<td>magnitude of static friction</td>
</tr>
<tr>
<td>$G$</td>
<td>gravitational constant</td>
</tr>
<tr>
<td>$G$</td>
<td>green quark color</td>
</tr>
<tr>
<td>$\bar{G}$</td>
<td>antigreen (magenta) antiquark color</td>
</tr>
<tr>
<td>$g$</td>
<td>acceleration due to gravity</td>
</tr>
<tr>
<td>$g$</td>
<td>gluons (carrier particles for strong nuclear force)</td>
</tr>
<tr>
<td>$h$</td>
<td>change in vertical position</td>
</tr>
<tr>
<td>$h$</td>
<td>height above some reference point</td>
</tr>
<tr>
<td>$h$</td>
<td>maximum height of a projectile</td>
</tr>
<tr>
<td>$h$</td>
<td>Planck's constant</td>
</tr>
<tr>
<td>$h$</td>
<td>height of the image</td>
</tr>
<tr>
<td>$h_o$</td>
<td>height of the object</td>
</tr>
<tr>
<td>$I$</td>
<td>electric current</td>
</tr>
<tr>
<td>$I$</td>
<td>intensity</td>
</tr>
<tr>
<td>$I$</td>
<td>intensity of a transmitted wave</td>
</tr>
<tr>
<td>$I$</td>
<td>moment of inertia (also called rotational inertia)</td>
</tr>
<tr>
<td>$I_0$</td>
<td>intensity of a polarized wave before passing through a filter</td>
</tr>
<tr>
<td>$I_{ave}$</td>
<td>average intensity for a continuous sinusoidal electromagnetic wave</td>
</tr>
<tr>
<td>$I_{rms}$</td>
<td>average current</td>
</tr>
<tr>
<td>$J$</td>
<td>joule</td>
</tr>
<tr>
<td>$J/\Psi$</td>
<td>Joules/psi meson</td>
</tr>
<tr>
<td>$K$</td>
<td>kelvin</td>
</tr>
<tr>
<td>$k$</td>
<td>Boltzmann constant</td>
</tr>
<tr>
<td>$k$</td>
<td>force constant of a spring</td>
</tr>
<tr>
<td>$K_\alpha$</td>
<td>x rays created when an electron falls into an $n = 1$ shell vacancy from the $n = 3$ shell</td>
</tr>
<tr>
<td>$K_\beta$</td>
<td>x rays created when an electron falls into an $n = 2$ shell vacancy from the $n = 3$ shell</td>
</tr>
<tr>
<td>kcal</td>
<td>kilocalorie</td>
</tr>
<tr>
<td>KE</td>
<td>translational kinetic energy</td>
</tr>
<tr>
<td>KE + PE</td>
<td>mechanical energy</td>
</tr>
<tr>
<td>KE$_e$</td>
<td>kinetic energy of an ejected electron</td>
</tr>
<tr>
<td>KE$_{rel}$</td>
<td>relativistic kinetic energy</td>
</tr>
<tr>
<td>KE$_{rot}$</td>
<td>rotational kinetic energy</td>
</tr>
<tr>
<td>KE</td>
<td>thermal energy</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>kg</td>
<td>kilogram (a fundamental SI unit of mass)</td>
</tr>
<tr>
<td>L</td>
<td>angular momentum</td>
</tr>
<tr>
<td>L</td>
<td>liter</td>
</tr>
<tr>
<td>L</td>
<td>magnitude of angular momentum</td>
</tr>
<tr>
<td>L</td>
<td>self-inductance</td>
</tr>
<tr>
<td>ℓ</td>
<td>angular momentum quantum number</td>
</tr>
<tr>
<td>L_{α\beta}</td>
<td>x rays created when an electron falls into an ( n = 2 ) shell from the ( n = 3 ) shell</td>
</tr>
<tr>
<td>L_e</td>
<td>electron total family number</td>
</tr>
<tr>
<td>L_μ</td>
<td>muon family total number</td>
</tr>
<tr>
<td>L_τ</td>
<td>tau family total number</td>
</tr>
<tr>
<td>L_f</td>
<td>heat of fusion</td>
</tr>
<tr>
<td>L_f \text{ and } L_v</td>
<td>latent heat coefficients</td>
</tr>
<tr>
<td>L_{orb}</td>
<td>orbital angular momentum</td>
</tr>
<tr>
<td>L_s</td>
<td>heat of sublimation</td>
</tr>
<tr>
<td>L_v</td>
<td>heat of vaporization</td>
</tr>
<tr>
<td>L_z</td>
<td>( z )- component of the angular momentum</td>
</tr>
<tr>
<td>M</td>
<td>angular magnification</td>
</tr>
<tr>
<td>M</td>
<td>mutual inductance</td>
</tr>
<tr>
<td>m</td>
<td>indicates metastable state</td>
</tr>
<tr>
<td>m</td>
<td>magnification</td>
</tr>
<tr>
<td>m</td>
<td>mass</td>
</tr>
<tr>
<td>m</td>
<td>mass of an object as measured by a person at rest relative to the object</td>
</tr>
<tr>
<td>m</td>
<td>meter (a fundamental SI unit of length)</td>
</tr>
<tr>
<td>m</td>
<td>order of interference</td>
</tr>
<tr>
<td>m</td>
<td>overall magnification (product of the individual magnifications)</td>
</tr>
<tr>
<td>m(A \text{X})</td>
<td>atomic mass of a nuclide</td>
</tr>
<tr>
<td>MA</td>
<td>mechanical advantage</td>
</tr>
<tr>
<td>m_e</td>
<td>magnification of the eyepiece</td>
</tr>
<tr>
<td>m_e</td>
<td>mass of the electron</td>
</tr>
<tr>
<td>m_ℓ</td>
<td>angular momentum projection quantum number</td>
</tr>
<tr>
<td>m_n</td>
<td>mass of a neutron</td>
</tr>
<tr>
<td>m_o</td>
<td>magnification of the objective lens</td>
</tr>
<tr>
<td>mol</td>
<td>mole</td>
</tr>
<tr>
<td>m_p</td>
<td>mass of a proton</td>
</tr>
<tr>
<td>m_s</td>
<td>spin projection quantum number</td>
</tr>
<tr>
<td>N</td>
<td>magnitude of the normal force</td>
</tr>
<tr>
<td>N</td>
<td>newton</td>
</tr>
<tr>
<td>N</td>
<td>normal force</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$N$</td>
<td>number of neutrons</td>
</tr>
<tr>
<td>$n$</td>
<td>index of refraction</td>
</tr>
<tr>
<td>$n$</td>
<td>number of free charges per unit volume</td>
</tr>
<tr>
<td>$N_A$</td>
<td>Avogadro's number</td>
</tr>
<tr>
<td>$N_r$</td>
<td>Reynolds number</td>
</tr>
<tr>
<td>$N \cdot m$</td>
<td>newton-meter (work-energy unit)</td>
</tr>
<tr>
<td>$N \cdot m$</td>
<td>newtons times meters (SI unit of torque)</td>
</tr>
<tr>
<td>OE</td>
<td>other energy</td>
</tr>
<tr>
<td>$P$</td>
<td>power</td>
</tr>
<tr>
<td>$P$</td>
<td>power of a lens</td>
</tr>
<tr>
<td>$P$</td>
<td>pressure</td>
</tr>
<tr>
<td>$p$</td>
<td>momentum</td>
</tr>
<tr>
<td>$p$</td>
<td>momentum magnitude</td>
</tr>
<tr>
<td>$p$</td>
<td>relativistic momentum</td>
</tr>
<tr>
<td>$P_{\text{tot}}$</td>
<td>total momentum</td>
</tr>
<tr>
<td>$P_{\text{tot}}'$</td>
<td>total momentum some time later</td>
</tr>
<tr>
<td>$P_{\text{abs}}$</td>
<td>absolute pressure</td>
</tr>
<tr>
<td>$P_{\text{atm}}$</td>
<td>atmospheric pressure</td>
</tr>
<tr>
<td>$P_{\text{atm}}$</td>
<td>standard atmospheric pressure</td>
</tr>
<tr>
<td>PE</td>
<td>potential energy</td>
</tr>
<tr>
<td>$PE_{\text{el}}$</td>
<td>elastic potential energy</td>
</tr>
<tr>
<td>$PE_{\text{elec}}$</td>
<td>electric potential energy</td>
</tr>
<tr>
<td>$PE_s$</td>
<td>potential energy of a spring</td>
</tr>
<tr>
<td>$P_g$</td>
<td>gauge pressure</td>
</tr>
<tr>
<td>$P_{\text{in}}$</td>
<td>power consumption or input</td>
</tr>
<tr>
<td>$P_{\text{out}}$</td>
<td>useful power output going into useful work or a desired, form of energy</td>
</tr>
<tr>
<td>$Q$</td>
<td>latent heat</td>
</tr>
<tr>
<td>$Q$</td>
<td>net heat transferred into a system</td>
</tr>
<tr>
<td>$Q$</td>
<td>flow rate—volume per unit time flowing past a point</td>
</tr>
<tr>
<td>$+Q$</td>
<td>positive charge</td>
</tr>
<tr>
<td>$-Q$</td>
<td>negative charge</td>
</tr>
<tr>
<td>$q$</td>
<td>electron charge</td>
</tr>
<tr>
<td>$q_p$</td>
<td>charge of a proton</td>
</tr>
<tr>
<td>$q$</td>
<td>test charge</td>
</tr>
<tr>
<td>QF</td>
<td>quality factor</td>
</tr>
<tr>
<td>$R$</td>
<td>activity, the rate of decay</td>
</tr>
<tr>
<td>$R$</td>
<td>radius of curvature of a spherical mirror</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>R</td>
<td>red quark color</td>
</tr>
<tr>
<td>_R</td>
<td>antired (cyan) quark color</td>
</tr>
<tr>
<td>R</td>
<td>resistance</td>
</tr>
<tr>
<td>R</td>
<td>resultant or total displacement</td>
</tr>
<tr>
<td>R</td>
<td>Rydberg constant</td>
</tr>
<tr>
<td>R</td>
<td>universal gas constant</td>
</tr>
<tr>
<td>r</td>
<td>distance from pivot point to the point where a force is applied</td>
</tr>
<tr>
<td>r</td>
<td>internal resistance</td>
</tr>
<tr>
<td>r⊥</td>
<td>perpendicular lever arm</td>
</tr>
<tr>
<td>r</td>
<td>radius of a nucleus</td>
</tr>
<tr>
<td>r</td>
<td>radius of curvature</td>
</tr>
<tr>
<td>r</td>
<td>resistivity</td>
</tr>
<tr>
<td>r or rad</td>
<td>radiation dose unit</td>
</tr>
<tr>
<td>rem</td>
<td>roentgen equivalent man</td>
</tr>
<tr>
<td>rad</td>
<td>radian</td>
</tr>
<tr>
<td>RBE</td>
<td>relative biological effectiveness</td>
</tr>
<tr>
<td>RC</td>
<td>resistor and capacitor circuit</td>
</tr>
<tr>
<td>rms</td>
<td>root mean square</td>
</tr>
<tr>
<td>rₙ</td>
<td>radius of the nth H-atom orbit</td>
</tr>
<tr>
<td>Rp</td>
<td>total resistance of a parallel connection</td>
</tr>
<tr>
<td>Rs</td>
<td>total resistance of a series connection</td>
</tr>
<tr>
<td>Rs</td>
<td>Schwarzschild radius</td>
</tr>
<tr>
<td>S</td>
<td>entropy</td>
</tr>
<tr>
<td>S</td>
<td>intrinsic spin (intrinsic angular momentum)</td>
</tr>
<tr>
<td>S</td>
<td>magnitude of the intrinsic (internal) spin angular momentum</td>
</tr>
<tr>
<td>S</td>
<td>shear modulus</td>
</tr>
<tr>
<td>S</td>
<td>strangeness quantum number</td>
</tr>
<tr>
<td>s</td>
<td>quark flavor strange</td>
</tr>
<tr>
<td>s</td>
<td>second (fundamental SI unit of time)</td>
</tr>
<tr>
<td>s</td>
<td>spin quantum number</td>
</tr>
<tr>
<td>s</td>
<td>total displacement</td>
</tr>
<tr>
<td>sec θ</td>
<td>secant</td>
</tr>
<tr>
<td>sin θ</td>
<td>sine</td>
</tr>
<tr>
<td>sz</td>
<td>z-component of spin angular momentum</td>
</tr>
<tr>
<td>T</td>
<td>period—time to complete one oscillation</td>
</tr>
<tr>
<td>T</td>
<td>temperature</td>
</tr>
<tr>
<td>Tc</td>
<td>critical temperature—temperature below which a material becomes a superconductor</td>
</tr>
<tr>
<td>T</td>
<td>tension</td>
</tr>
<tr>
<td>T</td>
<td>tesla (magnetic field strength θ)</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$t$</td>
<td>quark flavor top or truth</td>
</tr>
<tr>
<td>$t$</td>
<td>time</td>
</tr>
<tr>
<td>$t_{1/2}$</td>
<td>half-life—the time in which half of the original nuclei decay</td>
</tr>
<tr>
<td>tan $\theta$</td>
<td>tangent</td>
</tr>
<tr>
<td>$U$</td>
<td>internal energy</td>
</tr>
<tr>
<td>$\mu$</td>
<td>quark flavor up</td>
</tr>
<tr>
<td>$u$</td>
<td>unified atomic mass unit</td>
</tr>
<tr>
<td>$u$</td>
<td>velocity of an object relative to an observer</td>
</tr>
<tr>
<td>$u'$</td>
<td>velocity relative to another observer</td>
</tr>
<tr>
<td>$V$</td>
<td>electric potential</td>
</tr>
<tr>
<td>$V$</td>
<td>terminal voltage</td>
</tr>
<tr>
<td>$V$</td>
<td>volt (unit)</td>
</tr>
<tr>
<td>$V$</td>
<td>volume</td>
</tr>
<tr>
<td>$v$</td>
<td>relative velocity between two observers</td>
</tr>
<tr>
<td>$v$</td>
<td>speed of light in a material</td>
</tr>
<tr>
<td>$v$</td>
<td>velocity</td>
</tr>
<tr>
<td>$\bar{v}$</td>
<td>average fluid velocity</td>
</tr>
<tr>
<td>$V_B - V_A$</td>
<td>change in potential</td>
</tr>
<tr>
<td>$v_d$</td>
<td>drift velocity</td>
</tr>
<tr>
<td>$V_p$</td>
<td>transformer input voltage</td>
</tr>
<tr>
<td>$V_{rms}$</td>
<td>rms voltage</td>
</tr>
<tr>
<td>$V_s$</td>
<td>transformer output voltage</td>
</tr>
<tr>
<td>$v_{tot}$</td>
<td>total velocity</td>
</tr>
<tr>
<td>$v_w$</td>
<td>propagation speed of sound or other wave</td>
</tr>
<tr>
<td>$v_w$</td>
<td>wave velocity</td>
</tr>
<tr>
<td>$W$</td>
<td>work</td>
</tr>
<tr>
<td>$W$</td>
<td>net work done by a system</td>
</tr>
<tr>
<td>$W$</td>
<td>watt</td>
</tr>
<tr>
<td>$w$</td>
<td>weight</td>
</tr>
<tr>
<td>$w_{fl}$</td>
<td>weight of the fluid displaced by an object</td>
</tr>
<tr>
<td>$W_c$</td>
<td>total work done by all conservative forces</td>
</tr>
<tr>
<td>$W_{nc}$</td>
<td>total work done by all nonconservative forces</td>
</tr>
<tr>
<td>$W_{out}$</td>
<td>useful work output</td>
</tr>
<tr>
<td>$X$</td>
<td>amplitude</td>
</tr>
<tr>
<td>$^Z X_N$</td>
<td>notation for a particular nuclide</td>
</tr>
<tr>
<td>$x$</td>
<td>deformation or displacement from equilibrium</td>
</tr>
<tr>
<td>$x$</td>
<td>displacement of a spring from its undeformed position</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>x</td>
<td>horizontal axis</td>
</tr>
<tr>
<td>$X_C$</td>
<td>capacitive reactance</td>
</tr>
<tr>
<td>$X_L$</td>
<td>inductive reactance</td>
</tr>
<tr>
<td>$x_{rms}$</td>
<td>root mean square diffusion distance</td>
</tr>
<tr>
<td>y</td>
<td>vertical axis</td>
</tr>
<tr>
<td>$Y$</td>
<td>elastic modulus or Young's modulus</td>
</tr>
<tr>
<td>Z</td>
<td>atomic number (number of protons in a nucleus)</td>
</tr>
<tr>
<td>Z</td>
<td>impedance</td>
</tr>
</tbody>
</table>
Gluons, 1202
grand unified theory, 1204
Grand Unified Theory (GUT), 1201
gravitational constant, 205
gravitational constant, G, 214
gravitational potential energy, 230, 253
Gravitational waves, 1221
Gravitational waves, 1231
gray (Gy), 1154, 1175
greenhouse effect, 495, 497
grounded, 651, 654
grounding, 675, 688
GUT epoch, 1218, 1231
H
Hadrons, 1190
hadrons, 1205
half-life, 1129, 1140
Hall effect, 787, 801
Hall emf, 787, 801
harmonics, 607, 622
head, 89
head (of a vector), 114
head-to-tail method, 89, 115, 114
Hearing, 592, 611
hearing, 622
heat, 472, 497
heat engine, 512, 542
heat of sublimation, 483, 497
heat pump, 542
heat pump's coefficient of performance, 529
heat transfer, 432
Heisenberg uncertainty principle, 1051
Heisenberg's uncertainty principle, 1052, 1055
henny, 836, 850
hertz, 880
Higgs boson, 1203, 1204
high dose, 1155, 1175
holography, 1087
holography, 1103
Hooke's law, 175, 183
horizontally polarized, 978, 988
Hormes, 1156
hormesis, 1175
horsepower, 246, 253
Hubble constant, 1214, 1231
hues, 937, 949
Human metabolism, 511
human metabolism, 542
Huygens's principle, 957, 988
Hydrogen spectrum wavelength, 1073
hydrogen spectrum wavelengths, 1103
hydrogen-like atom, 1074, 1103
hydrogen-spectrum wavelengths, 1072
hyperopia, 933, 949
inductive reactance, 842, 850
inductor, 837, 850
inelastic collision, 273, 282
inertia, 128, 155
Inertia, 156
inertial confinement, 1165, 1175
inertial frame of reference, 144, 155, 998, 1021
inflationary scenario, 1218, 1231
Infrared radiation, 671
infrared radiation, 1055
infrared radiation (IR), 880
Infrared radiation, 1040
infrasound, 611, 622
ink jet printer, 651
ink-jet printer, 655
Instantaneous acceleration, 45
Instantaneous acceleration, 75
Instantaneous speed, 41
Instantaneous speed, 75
Instantaneous velocity, 40
Instantaneous velocity, 76
insulator, 651
insulators, 626
intensity, 580, 582, 597, 622, 878, 880
intensity reflection coefficient, 617, 622
Interference microscopes, 896
interference microscopes, 988
internal energy, 509, 542
Internal kinetic energy, 271
internal kinetic energy, 282
internal resistance, 744, 765
intraocular pressure, 388, 391
intrinsic magnetic field, 1091, 1103
intrinsic spin, 1091, 1103
ionizing radiation, 1056, 1055, 1116, 1140
ionosphere, 648, 665
irreversible process, 519, 542
isobaric process, 514, 542
isochoric, 515
isochoric process, 542
isolated system, 269, 282
isothermal, 517
isothermal process, 542
isotopes, 1121, 1141
J
joule, 226, 254
Joule's law, 737, 765
junction rule, 731, 765
K
Kelvin, 433
Kelvin scale, 464
kilocalorie, 472, 497
kilogram, 20, 31
kilowatt-hour, 254
kilowatt-hours, 248
kinematics, 36, 76, 102, 114
kinematics of rotational motion, 324, 349
kinetic energy, 228, 254
kinetic friction, 166, 183
Kirchhoff's rules, 750, 765
L
Laminar, 409
laminar, 421
laser, 1084, 1103
laser printer, 655
Laser printers, 651
Laser vision correction, 936
laser vision correction, 949
latent heat coefficient, 497
latent heat coefficients, 479
law, 15, 31
law of conservation of angular momentum, 341, 349
law of conservation of charge, 634, 655
law of conservation of energy, 242, 254
law of inertia, 128, 156, 155
law of reflection, 921
law of refraction, 894
Length contraction, 1006
length contraction, 1021
Lenz's law, 817, 850
leptons, 1190, 1205
linear accelerator, 1189, 1205
linear hypothesis, 1155, 1175
Linear momentum, 264
linear momentum, 282
liquid drop model, 1167, 1175
liter, 400, 421
longitudinal wave, 574, 582
loop rule, 751, 765
Lorentz force, 782, 801
loudness, 611, 622
low dose, 1155, 1175
M
MACHOs, 1225, 1231
macrostate, 539, 542
magic numbers, 1123, 1141
magnetic confinement, 1164, 1175
magnetic damping, 822, 850
magnetic field, 781, 802, 864, 880
magnetic field lines, 781, 801, 880
Magnetic field lines, 863
magnetic field strength, 880
magnetic field strength (magnitude) produced by a long straight current-carrying wire, 795, 802
magnetic field strength at the center of a circular loop, 796, 802
magnetic field strength inside a solenoid, 797, 802
magnetic flux, 816, 850
magnetic force, 782, 802
magnetic monopoles, 780, 802
Magnetic resonance imaging (MRI), 800
magnetic resonance imaging (MRI), 802
magnitized, 778, 802
magnetocardiogram (MCG), 801, 802
magnetoeencephalogram (MEG), 801, 802
magnification, 909, 921
magnitude, 90
magnitude (of a vector), 114
magnitude of kinetic friction, 183
magnitude of kinetic friction, $f_k$, 167
magnitude of static friction, 183
magnitude of static friction, $f_s$, 166
magnitude of the intrinsic (internal) spin angular momentum, 1094, 1103
mass, 128, 155
Mass, 156
mass number, 1120, 1141
massive compact halo objects, 1225
maximum field strength, 878, 880
Maxwell's equations, 796, 801, 862, 880
mechanical advantage, 303, 311
mechanical energy, 236, 254, 688
Mechanical energy, 669
mechanical equivalent of heat, 472, 497
meson, 1185, 1205
Mesons, 1193
metabolic rate, 249, 254
rainbow, 921
range, 107, 115
range of radiation, 1116, 1141
rate of conductive heat transfer, 485, 497
rate of decay, 1131, 1141
ray, 888, 921
Ray tracing, 906
Rayleigh criterion, 970, 989
RC circuit, 765
real image, 908, 921
reflected light is completely polarized, 981
reflected light that is completely polarized, 989
refraction, 891, 921
relative biological effectiveness, 1154
relative biological effectiveness (RBE), 1175
relative humidity, 460, 464
relative osmotic pressure, 421, 422
relative velocities, 112
relative velocity, 115
relativistic Doppler effects, 1023, 1021
Relativistic kinetic energy, 1018
relativistic kinetic energy, 1022
Relativistic momentum, 1014
relativistic momentum, 1022
relativistic velocity addition, 1010, 1022
Relativity, 18
relativity, 31, 112, 115, 998, 1022
Renewable forms of energy, 251
renewable forms of energy, 254
resistance, 703, 725, 736, 765
resistivity, 705, 725
resistor, 736, 761, 765
resonance, 571, 582
resonant, 865, 880
resonant frequency, 846, 850
resonate, 731, 582
Rest energy, 1015
rest energy, 1022
rest mass, 1014, 1022
restoring force, 552, 582
resultant, 90, 115
resultant vector, 89, 115
retinex, 949
retinex theory of color vision, 939, 949
retinexes, 939
reverse dialysis, 421, 422
Reverse osmosis, 421
reverse osmosis, 422
reversible process, 518, 542
Reynolds number, 416, 422
right hand rule 1, 782
right hand rule 1 (RHR-1), 802
right hand rule 2, 795
right hand rule 2 (RHR-2), 802
right-hand rule, 346, 349
RLC circuit, 880
rms current, 713, 725
rms voltage, 713, 725
rods and cones, 957, 949
roentgen equivalent man, 1154
roentgen equivalent man (rem), 1175
rotation angle, 190, 214
rotational inertia, 328, 349
rotational kinetic energy, 331, 349
Rydberg constant, 1072, 1077, 1103
S
saturation, 460, 464
scalar, 38, 76, 94, 115, 670, 689
Schwarzschild radius, 1220, 1231
scientific method, 16, 31
scintillators, 1118, 1141
screening, 646, 655
second, 19, 31
second law of motion, 264, 283
second law of thermodynamics, 520, 521, 524, 543
second law of thermodynamics stated in terms of entropy, 543
second postulate of special relativity, 999, 1022
Self-inductance, 837
self-inductance, 850
semipermeable, 420, 422, 720, 725
series, 736, 765
shear deformation, 180, 183
shell, 1098, 1103
shielding, 1156, 1175
shock hazard, 716, 725, 832, 851
short circuit, 716, 725
shunt resistance, 756, 765
SI unit of torque, 295
SI units, 19, 31
SI units of torque, 311
siervert, 1154, 1175
significant figures, 27, 31
simple circuit, 703, 725
Simple Harmonic Motion, 557
simple harmonic motion, 582
simple harmonic oscillator, 557, 582
simple pendulum, 561, 582
simplified theory of color vision, 937, 949
single-photon-emission computed tomography (SPECT), 1175
single-photon-emission computed tomography(SPECT), 1152
slope, 69, 76
solenoid, 796, 802
Solid-state radiation detectors, 1119
solid-state radiation detectors, 1141
sonic boom, 603, 622
sound, 592, 622
sound intensity level, 598, 622
sound pressure level, 600, 622
south magnetic pole, 776, 802
space quantization, 1091, 1103
special relativity, 1022
special relativity, 998
specific gravity, 376, 391
specific heat, 474, 497
speed of light, 880
spin projection quantum number, 1095, 1103
spin quantum number, 1096, 1103
spontaneous symmetry breaking, 1218, 1231
stable equilibrium, 297, 311
Standard Model, 1202
standard model, 1205
standing wave, 577, 865, 880
static electricity, 630, 655
static equilibrium, 292, 300, 311
static friction, 166, 183
statistical analysis, 540, 543
Stefan-Boltzmann law of radiation, 494, 497
step-down transformer, 830, 851
step-up transformer, 830, 851
stimulated emission, 1084
stimulated emission, 1103
Stokes’ law, 174, 183
strain, 180, 183
strong, 1195, 1205
strangeness, 1193, 1205
stress, 180, 183
sublimation, 458, 464, 497
Sublimation, 482
subshell, 1098, 1103
Superconductors, 1227, 1231
supercriticality, 1169, 1175
superforce, 1218, 1231
superposition, 575, 582
Superstring theory, 1203, 1223, 1231
superstring theory, 1205
surface tension, 380, 391
synchrontron, 1187, 1205
synchrontron radiation, 1188, 1205
system, 128, 156
systolic pressure, 371, 391
Systolic pressure, 387
T
tagged, 1150, 1175
tail, 89, 115
tangential acceleration, 321, 349
tau family number, 1205
Television, 869
Temperature, 432
temperature, 464
temperature coefficient of resistivity, 707, 725
tensile strength, 177, 183
tension, 139, 156
terminal speed, 417, 422
terminal voltage, 746, 765
tesla, 782, 802
test charge , 641, 655
the second law of thermodynamics stated in terms of entropy, 534
theory, 15, 31
theory of quark confinement, 1200, 1205
therapeutic ratio, 1159, 1175
thermal agitation, 871, 880
thermal conductivity, 485, 497
thermal energy, 238, 242, 254, 452, 464
thermal equilibrium, 437, 464
thermal expansion, 438, 464
thermal hazard, 716, 725, 832, 851
thermal stress, 442
thermal stress, 464
thin film interference, 974, 989
thin lens, 906
thin lens equations, 909
thought experiment, 1218, 1231
three-wire system, 832, 851
thrust, 134, 156, 156
timbre, 611, 622
time, 39, 76
Time dilation, 1001
time dilation, 1022
TOE epoch, 1218, 1231
tone, 611, 622
top, 1200, 1205
Torque, 294
torque, 311, 328, 349
Total energy, 1015
total energy, 1022
total internal reflection, 895
trajectory, 101, 115
transformer, 851
transformer equation, 830, 851
Transformers, 828
transverse wave, 574, 582, 865, 880
triple point, 458, 464
Tunneling, 1139
tunneling, 1141
turbulence, 409, 422
TV, 880
twin paradox, 1022
U
ultra high frequency, 869
ultra-high frequency (UHF), 880
ultracentrifuge, 195, 214
ultrasound, 611, 622
ultraviolet (UV) microscopes, 985
ultraviolet (UV) microscopes, 989
ultraviolet radiation, 1038
ultraviolet radiation, 1056
ultraviolet radiation (UV), 873, 880
uncertainty, 26, 31
uncertainty in energy, 1052, 1056
uncertainty in momentum, 1051, 1056
uncertainty in position, 1051, 1056
uncertainty in time, 1052, 1056
under damping, 582
underdamped, 568
uniform circular motion, 190, 214
units, 18, 32
unpolarized, 978, 989
unstable equilibrium, 297, 311
up, 1195, 1205
useful work, 249, 254

V
Van de Graaff, 1187, 1205
Van de Graaff generator, 655
Van de Graaff generators, 650
vapor, 459, 464
Vapor pressure, 459
vapor pressure, 464
vector, 38, 76, 88, 115, 655, 670, 689
vector addition, 109, 115, 643, 655
vectors, 87, 642
velocity, 109, 115
vertically polarized, 978, 989
very high frequency, 869
very high frequency (VHF), 880
virtual image, 911, 921
virtual particles, 1184, 1205
viscosity, 410, 422
viscous drag, 417, 422
Visible light, 872
visible light, 880, 1039, 1056
voltage, 667, 736, 765
voltage drop, 736, 765
voltmeter, 765
Voltmeters, 754

W
watt, 246, 254
wave, 573, 582
wave velocity, 573, 582
wavelength, 573, 582, 864, 880
wavelength in a medium, 957, 989
waves, 552
weakly interacting massive particles, 1225
weight, 130, 156
Weight, 137
Wheatstone bridge, 760, 765
WIMPs, 1225, 1231
work, 224, 254
work-energy theorem, 228, 254, 333, 350, 349

X
x ray, 1056
x rays, 1037, 1103
X rays, 1077
X-ray, 875, 880
x-ray diffraction, 1080, 1103
xerography, 651, 655

Y
y-intercept, 69, 76