

# *Feynman Simplified*

## *3C: Quantum Mechanics*

### *Part Three*

Everyone's Guide to the  
**Feynman Lectures**

Robert L. Piccioni, Ph.D.

Real Science Publishing

**Feynman Simplified**  
**3C: Quantum Mechanics**  
**Part Three**

**Everyone's Guide**  
**to the**  
*Feynman Lectures on Physics*

**by**  
**Robert L. Piccioni, Ph.D.**

**Copyright © 2014**  
**by**  
**Robert L. Piccioni**

Published by  
Real Science Publishing  
3949 Freshwind Circle  
Westlake Village, CA 91361, USA

Edited by Joan Piccioni

All rights reserved, including the right of  
reproduction in whole or in part, in any form.

Visit our web site  
[www.guidetothecosmos.com](http://www.guidetothecosmos.com)

# Everyone's Guide to the Feynman Lectures on Physics

*Feynman Simplified* gives mere mortals access to the fabled *Feynman Lectures on Physics*.

Quantum mechanics is rarely taught well in introductory physics courses, largely because this challenging subject was not well taught to many of today's instructors. Few had the opportunity to learn quantum mechanics from some who understood it profoundly; almost none learned it from one of its creators. Here more than anywhere else, Feynman excels. Here more than anywhere else, *Feynman Simplified* can help you learn from the very best, but at a humane pace.

# This Book

*Feynman Simplified: 3A* covers the final third of Volume 3 of *The Feynman Lectures on Physics*. The topics we explore include:

- Quantum Behavior of Elementary Particles
- Angular Momentum & Rotations of Any Spin
- Electron Atomic States & The Periodic Table
- Philosophy of Wave Functions & Probability
- Macroscopic QM: Superconductivity
- Entanglement, Schrödinger's Cat & Teleportation
- EPR Paradox: QM vs. Local Realism
- Alternative Interpretations of QM

To find out about other eBooks in the *Feynman Simplified* series, click [HERE](#).

I welcome your comments and suggestions. Please contact me through my [WEBSITE](#).

If you enjoy this eBook please do me the great favor of rating it on Amazon.com or BN.com.

# Table of Contents

|                                                    |  |
|----------------------------------------------------|--|
| Review: Summary of Part One                        |  |
| Review: Summary of Part Two                        |  |
| Chapter 25: Particle Physics: $\Lambda^0$ Decay    |  |
| Chapter 26: Particle Physics: Kaons                |  |
| Chapter 27: Angular Momentum                       |  |
| Chapter 28: Rotation Of Any Spin                   |  |
| Chapter 29: Electron Wave Functions in Hydrogen    |  |
| Chapter 30: The Periodic Table of Elements         |  |
| Chapter 31: Quantum & Algebraic Operators          |  |
| Chapter 32: Probability & Wave Functions           |  |
| Chapter 33: Superconductivity                      |  |
| Chapter 34: Collapse, Entanglement & Teleportation |  |
| Chapter 35: EPR: QM vs. Local Realism              |  |
| Chapter 36: Review of QM                           |  |
| Feynman's Epilogue                                 |  |

# Summary of QM Part One

---

**Particle-wave duality:** Everything in our universe has, in varying degrees, the properties of both classical particles and classical waves. In particular, particles have wavelengths given by:  $\lambda = h/p$ .

**Uncertainty Principle:** the wave properties of particles preclude the possibility of simultaneously measuring both their position and momentum with unlimited precision, as quantified in four equations:

$$\Delta x \Delta p_x \geq \hbar/2$$

$$\Delta y \Delta p_y \geq \hbar/2$$

$$\Delta z \Delta p_z \geq \hbar/2$$

$$\Delta t \Delta E \geq \hbar/2$$

where  $\hbar = h/2\pi = 1.055 \times 10^{-34}$  joule-sec, and  $h = 6.626 \times 10^{-34}$  joule-sec is Planck's constant.

A **quantum state** defines all the variable properties that an entity can have, such as position, linear momentum, energy, spin, and angular momentum. It does not define intrinsic properties, such as charge and mass. In general, different types of particles can be put into a given state, and can be moved from one state to another. It is beneficial to think of a quantum state as a vector that defines a location in the space of all possible properties.

**Spin** is a quantized form of angular momentum that is an intrinsic and immutable property of each type of elementary particle. All fermions have half-integral spin:  $s = n/2$ , for some odd integer  $n > 0$ . All elementary fermions have spin  $s = 1/2$ , as do protons and neutrons, which are composite particles. Bosons have integral spin:  $s = 0, 1$ , or  $2$ .

There is an essential distinction between a particle's *spin* and its *component of spin* along any axis. Like other forms of angular momentum, spin is a vector with magnitude and direction. The spin of every particle never changes, although the orientation of its spin may change. The primary fermions, those with spin  $1/2$ , can have a component of spin along any selected axis of  $+\hbar/2$  or  $-\hbar/2$ , which are called *spin up* and *spin down*. The Z and W bosons with spin  $1$  can have a component of spin along any selected axis of  $+\hbar$ ,  $0$ , or  $-\hbar$ . In general, a particle of spin  $s$  can have a component of spin along any selected axis of  $-\hbar s, -(s-1)\hbar, \dots, +(s-1)\hbar, +\hbar s$ ; the allowed values of spin components are always separated by integral multiples of  $\hbar$ . Photons have spin  $1$ , but cannot have a zero component of spin along any axis, because they are massless. A particle with spin  $s$  has a spin vector whose magnitude equals  $\sqrt{s(s+1)}\hbar$ ; this magnitude is rarely mentioned, since it is so much easier to say "spin  $s$ ."

**Measurement** is a contact sport, more like boxing than dancing. Observation requires interaction. Substantial interactions alter the behavior of what is observed. Every measurement forces the observed entity into a definite state, one with an allowed value for that measurement. Particles have

no memory of their prior history. If an atom's spin is measured to be +1 along the z-axis, nothing else about its spin is knowable.

Feynman's **General Principles** of probabilistic quantum mechanics are:

**First:** the probability  $\text{Prob}(y)$  of event  $y$  is proportional to the square of the magnitude of probability amplitude  $\phi(y)$ , which is a complex number. If event  $y$  can occur in only one way,  $P(y)$  is:

$$\text{Prob}(y) = \phi\phi^* = |\phi|^2$$

In Dirac's bra-ket notation, the amplitude that A results in, or goes to, B is  $\langle B|A \rangle$ . This is analogous to the dot product of normal vectors A and B. For example, in the two-slit experiment:

$$\phi(y) = \langle \text{particle arrives at } y \mid \text{particle leaves source } S \rangle$$

$$\phi(y) = \langle y|S \rangle$$

**Second:** when event  $y$  can occur in  $N$  indistinguishable ways, the amplitude of  $y$  equals the sum of the amplitudes for each separate way:

$$\langle y|S \rangle = \sum_K \langle y|K \rangle \langle K|S \rangle, \text{ sum } K=1 \dots N$$

We sum amplitudes only when each separate path results in exactly the same final state.

As the two-slit experiment demonstrates, with no substantial disturbances, it is impossible — for both man and nature — to *distinguish* the paths by which particles travel to reach an event  $y$ . Paths are *distinguished* only when particle-waves along those paths are substantially disturbed. What matters is the magnitude of the disturbance, not whether it occurs naturally or is human-directed. But due to our scale, any human-directed interaction will almost certainly substantially disturb a quantum system.

**Third:** when event  $y$  can occur in  $N$  distinguished ways, the probability of  $y$  equals the sum of the probabilities for each separate way:

$$\text{Prob}(y) = \sum_K |\langle y|K \rangle \langle K|S \rangle|^2, \text{ sum } K=1 \dots N$$

**Fourth:** the amplitude for a sequence of events equals the product of the amplitudes for each separate event. The amplitude for a particle to go from  $S$  to  $x$  and then from  $x$  to  $y$  is:  $\langle y|x \rangle \langle x|S \rangle$ .

**Fifth:** For multiple particles, if  $\phi_1$  is the amplitude of particle 1 going from  $S_1$  to  $F_1$ , and  $\phi_2$  is the amplitude of particle 2 going from  $S_2$  to  $F_2$ , then the amplitude for both events to occur is  $\phi_1 \times \phi_2$ .

**Sixth:** When two intrinsically identical particles can enter, exit, or be in the same state, their amplitudes interfere. If the particles are bosons governed by **Bose-Einstein statistics**, their amplitudes add. If the particles are fermions governed by **Fermi-Dirac statistics**, their amplitudes subtract. There is no third alternative. For two identical particles 1 and 2, and any two indistinguishable states A and B, the combined amplitude is:

Fermions:  $\langle 1|A\rangle\langle 2|B\rangle - \langle 1|B\rangle\langle 2|A\rangle$

Bosons :  $\langle 1|A\rangle\langle 2|B\rangle + \langle 1|B\rangle\langle 2|A\rangle$

Particles with different spins are distinguishable.

All fundamental particles of each type are **intrinsically exactly identical**. While not fundamental, all protons are exactly identical to one another as are all neutrons.

The class of **fermions** includes protons, electrons, neutrons, neutrinos, and quarks. All particles that form material objects are fermions with spin 1/2.

The class of **bosons** includes photons, gluons,  $W^\pm$ ,  $Z^0$ , and the Higgs boson. All forces between elementary particles are carried by bosons. Photons are the force exchange particles of electromagnetism. Gluons mediate the strong force. The  $W^\pm$  and  $Z^0$  *intermediate vector bosons* mediate the weak force. See [\*Higgs & Bosons & Fermions...\*](#) to further explore elementary particle physics.

The probability that **N identical bosons** are in a common state is larger by a factor of  $N!$  than the probability of  $N$  non-identical particles being in that state.

The probability of one more boson entering a state occupied by  $N$  identical bosons is  $N+1$  times greater than it would be if the state were empty. Bosons are groupies: the bigger the party, the greater the attraction.

The gregarious behavior of bosons underlies Einstein's laws of radiation and the operation of lasers.

Fermions are antisocial; they never exist in the same state as another identical fermion. This is called the **Pauli Exclusion Principle**. Two fermions are not in the same state if their spins are different.

Nature has room for both groupies and individualists.

## Similarity of Vectors & States

In normal 3-D space, a vector  $\mathbf{v}$  specifies any selected point within the space. In quantum mechanics,  $|S\rangle$  specifies a *state*, any selected set of particle properties.

In 3-D, we choose a complete set of *orthonormal basis vectors*  $\mathbf{e}_x$ ,  $\mathbf{e}_y$  and  $\mathbf{e}_z$ .

In QM, we choose a complete set of *orthonormal basis states*  $|J\rangle$ ,  $J=1,\dots,N$ .

Orthonormal means for each  $J$  and  $K$ :  $\mathbf{e}_J \bullet \mathbf{e}_K = \delta_{JK}$  in 3D; and  $\langle K|J\rangle = \delta_{JK}$  in QM. Here  $\delta_{JK}$  is the *Kronecker delta*, which equals 1 if  $J=K$  and equals 0 otherwise.

In QM, a basis is “complete” if every possible state is some linear combination of basis states — for

any state  $|\varphi\rangle$ :

$$|\varphi\rangle = \sum_j a_j |J\rangle = a_1 |1\rangle + a_2 |2\rangle + \dots a_N |N\rangle$$

$$\langle\varphi| = \sum_j \langle J|\varphi\rangle \langle J|$$

where  $a_j = \langle J|\varphi\rangle$  are *complex* numbers. The set of all  $a_j$  completely defines  $|\varphi\rangle$ .

## Differences Between Vectors & States

In 3-D, the dot product is commutative:  $\mathbf{A}\cdot\mathbf{B} = \mathbf{B}\cdot\mathbf{A}$ . But, in QM the product of two states is **not** commutative:  $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$ , which is the complex conjugate of  $\langle\psi|\varphi\rangle$ . The complex conjugate of a complex quantity is obtained by replacing  $i$  with  $-i$ . In general  $\langle\psi|\varphi\rangle$  and  $\langle\varphi|\psi\rangle$  are not equal.

Since  $\langle\varphi|\psi\rangle = \sum_j \langle\varphi|J\rangle \langle J|\psi\rangle$  for all states  $\varphi$  and  $\psi$ , Feynman writes the **open equation**:

$$|\varphi\rangle = \sum_j |J\rangle \langle J|\varphi\rangle$$

Open equations are valid when multiplied by any bra and/or by any ket.

## Other Bra-Ket Notes

Bras and kets are essential intermediate steps in solving problems, but they are not end results by themselves. Calculations of measurable quantities must not end with expressions having unmatched bras or kets.

Bra and ket are complex conjugates:  $\langle\psi| = |\psi\rangle^*$ . This means  $\langle\psi|$  and  $|\psi\rangle$  are different, in general. Either could properly represent a quantum state. Feynman says he always uses kets  $|\psi\rangle$  to identify states, for consistency.

**Operators** transform one state vector into another. Any action that changes particle states is represented in quantum mechanics by an operator. As an example:

$$\langle\chi|A|\psi\rangle = \sum_{jk} \langle\chi|K\rangle \langle K|A|J\rangle \langle J|\psi\rangle$$

denotes the product of state  $\chi$  with the vector resulting from  $A$  operating on state  $\psi$ , which is shown expanded in basis states. Knowing  $\langle K|A|J\rangle$  for all basis states  $J$  and  $K$  completely determines  $A$ .

Operators are often represented as  $N\times N$  matrices, where  $N$  is the number of basis states. Matrix component  $A_{kj}$  equals  $\langle K|A|J\rangle$ . If multiple operators act on a state sequentially, we express that using the product of those operators. Matrix products are not commutative, so the order of factors is essential. The first operator to act holds the right-most position, with the second immediately on its left, and continuing on to the last operator to act in the left-most position. For two operators,  $A$  followed by  $B$ , the expression is:

$$\langle \chi | B A | \psi \rangle = \sum_{jkl} \langle \chi | L \rangle \langle L | B | K \rangle \langle K | A | J \rangle \langle J | \psi \rangle$$

**Waves in an Energy Potential:** the probability amplitude for a particle of momentum  $p$  and energy  $E$  (kinetic plus mass), with potential energy  $V$ , is proportional to:

$$\exp \{ (E+V)t/i\hbar - \mathbf{p} \cdot \mathbf{r} / i\hbar \}$$

The **Hamiltonian equation** governs the time evolution of states:

$$i\hbar d|\Psi(t)\rangle/dt = H |\Psi(t)\rangle$$

## Review of Two-State Systems

For a system with two basis states  $|1\rangle$  and  $|2\rangle$ , any state  $|\Psi\rangle$  can be expressed as a linear combination of the basis states:

$$|\Psi\rangle = |1\rangle \langle 1|\Psi\rangle + |2\rangle \langle 2|\Psi\rangle$$

$$|\Psi\rangle = C_1 |1\rangle + C_2 |2\rangle$$

The Hamiltonian equation yields two independent differential equations:

$$i\hbar d(C_1)/dt = H_{11} C_1 + H_{12} C_2$$

$$i\hbar d(C_2)/dt = H_{21} C_1 + H_{22} C_2$$

We can also write this more compactly:

$$i\hbar dC_j/dt = \sum_k H_{jk} C_k, \text{ for any } j, \text{ sum over } k$$

Here  $H_{12}$  is the amplitude for state  $|2\rangle$  to transition to state  $|1\rangle$ , and  $H_{21}$  is the amplitude for state  $|1\rangle$  to transition to state  $|2\rangle$ .

Assuming the Hamiltonian components are constant in time, the Hamiltonian equation has two solutions that are stationary states, states of definite energy. These are:

$$|\psi_+\rangle = |+\rangle \exp\{E_+ t/i\hbar\}, \text{ with } E_+ = E_0 + E^*$$

$$|\psi_-\rangle = |-\rangle \exp\{E_- t/i\hbar\}, \text{ with } E_- = E_0 - E^*$$

where,

$$E_0 = (H_{11} + H_{22})/2$$

$$E^* = \sqrt{[(H_{11} - H_{22})^2/4 + H_{12} H_{21}]}$$

$$|+\rangle = a_{1+} |1\rangle + a_{2+} |2\rangle$$

$$|-\rangle = a_{1-} |1\rangle + a_{2-} |2\rangle$$

In V3p10-2, Feynman says the coefficients of the basis states, the a's in the above equation, are complex constants that satisfy:

$$1 = |a_{1+}|^2 + |a_{2+}|^2$$

$$1 = |a_{1-}|^2 + |a_{2-}|^2$$

$$a_{1+} / a_{2+} = H_{12} / (E_+ - H_{11})$$

$$a_{1-} / a_{2-} = H_{12} / (E_- - H_{11})$$

In the special case of a symmetric two-state system, where  $H_{11}=H_{22}$  and  $H_{12}=H_{21}=E^*$ :

$$a_{1+} / a_{2+} = -E^* / (E_+ - E_0) = -1$$

$$a_{1-} / a_{2-} = -E^* / (E_- - E_0) = +1$$

$$|+\rangle = \{ |1\rangle - |2\rangle \} / \sqrt{2}$$

$$|-\rangle = \{ |1\rangle + |2\rangle \} / \sqrt{2}$$

# Summary of QM Part Two

---

## Particle Exchange Model of Force

Quantum Field Theory (QFT) states that the electromagnetic, strong, and weak forces arise from fermions (particles of matter) exchanging **virtual** particles (most commonly bosons). **Virtual** particles are not directly observable and may have exotic properties, including negative energy and imaginary momentum. Gravity has not been successfully incorporated within the QFT model. The force-exchange bosons are:

- Photons for the electromagnetic force
- $W^+$ ,  $W^-$ ,  $Z^0$  for the weak force
- Gluons for the strong force

## Barrier Penetration

A particle's probability amplitude  $A$  declines exponentially within a barrier in which its kinetic energy is negative and its momentum  $p$  is imaginary, according to:

$$A \sim \exp\left\{\frac{i}{\hbar} \int p \, dz\right\}$$

Here  $z$  is the distance inside the barrier. What penetrates the barrier is the particle's probability amplitude, something that has no physical reality. A particle on one side of a classically impenetrable barrier has a small but non-zero probability of appearing later on the opposite side of the barrier. But quantum mechanics says no real particle can ever be observed to have negative energy, negative mass, or imaginary momentum; hence particles can never be observed within such barriers.

**Matrices** are arrays of components laid out in rows and columns. In quantum mechanics, the components are typically amplitudes or numbers. An  $n \times m$  ("n-by-m") matrix has  $n$  rows,  $m$  columns, and  $n \cdot m$  components. For matrices,  $A$ ,  $B$ , and  $C$ , and scalar  $s$ , basic matrix arithmetic includes:

1. Multiplication by scalar  $s$ :  $A_{ij} = sB_{ij}$
2. Matrix addition:  $A_{ij} = B_{ij} + C_{ij}$
3. Matrix multiplication:  $C_{ij} = \sum_k A_{ik} B_{kj}$
4. Unit matrix  $\delta_{ij}$  has components: 1 if  $i=j$ , else zero
5. If  $A^{-1}$  is the inverse of  $A$ ,  $A^{-1}A = AA^{-1} = \delta_{ij}$
6. Only non-singular  $n \times n$  matrices have inverses.

The general equation for the determinant of an N×N matrix M is:

$$\text{Det } |M_{ij}| = \sum \text{Sign}(abc\dots) \cdot M_{1a} \cdot M_{2b} \cdot M_{3c} \cdot \dots$$

Here, abc... is a permutation of the integers 1 through N. The sum extends over all permutations, and Sign(abc..) equals +1 for even permutations and −1 for odd permutations.

In a **magnetic field**, the Hamiltonian equation for an electron can be written in terms of Pauli spin matrices:

$$i\hbar \frac{d}{dt} |\Psi\rangle = -\mu \boldsymbol{\sigma} \cdot \mathbf{B} |\Psi\rangle$$

**Electron atomic orbital states** are characterized by three integer quantum numbers: *principal quantum number* n; *angular momentum quantum number* j; and *magnetic quantum number* m, which is the component of j along any selected axis. Modern convention employs lower case l rather than j, but I follow Feynman's notation to enhance eBook readability. Each orbital electron state is further subdivided into two spin states.

## Electrons in Crystals

Feynman says an electron added to an array of identical, electrically neutral, equally spaced, immovable atoms “can ride right through the crystal and flow perfectly freely even though it has to hit all the atoms [in its path]. ... That is how a solid can conduct electricity.”

In a one-dimensional crystal, the Hamiltonian equations for an added electron are:

$$i\hbar \frac{dC_n}{dt} = E_0 C_n - A C_{n-1} - A C_{n+1}$$

Here  $E_0$  is the zero-point energy of a stationary electron;  $C_n$  is the amplitude for the electron to be at atom #n; and A is the amplitude for the electron to transition to an adjacent atom. For atomic spacing b, the stationary states are given by:

$$C_n = \exp\{i(knb - Et/\hbar)\}$$

$$E(k) = E_0 - 2A \cos(kb)$$

The electron's energy is in a **conduction band** between  $E_0 - 2A$  and  $E_0 + 2A$ , with wave number k between  $-\pi/b$  and  $+\pi/b$ . Interaction with matter changes the electron's **effective mass** to:

$$m_{\text{eff}} = \hbar^2 / (2Ab^2)$$

The same mathematics applies to the propagation of other entities through crystals, including:

- An electron in an *excited state*.
- A missing electron, called a *hole*.

- Added energy, called an *exciton*.

An **N-atom ring molecule** is analyzed as an infinite line of atoms that repeat every N atoms. The eigenstates are (for integer m):

$$\begin{aligned} a_n(k) &= \exp\{iknb\}, \text{ for } n=1 \dots N \\ k_m &= (2\pi/Nb)m, \text{ for } -N/2 < m \leq +N/2 \\ E &= E_0 - 2A \cos(2\pi m/N) \end{aligned}$$

An **N-atom linear molecule** is analyzed as an infinite line of atoms with zero amplitudes at atom n=0 and n=N+1. The eigenstates are (for integer m):

$$\begin{aligned} a_n(k) &= \sin(knb), \text{ for } n=1 \dots N \\ k &= \pi m/[b(N+1)], \text{ for } m=1 \dots N \\ E &= E_0 - 2A \cos(kb) \end{aligned}$$

**Semiconductors** are good insulators at low temperatures, and are modestly conducting at room temperature. Electric current in a semiconductor is carried by both **negative carriers** (free electrons) and **positive carriers** (holes, vacant electron orbits in crystal atoms). The minimum energy required to create a negative/positive carrier pair within a semiconductor is its **band gap**,  $E_{\text{gap}}$ . Electron-hole pairs can be created by: photon absorption in solar cells; ionization by high-energy particles; and conversion of thermal energy. In a pure crystal, negative and positive carrier densities are equal. Adding **dopants** increases one carrier density while reducing the other; their product is always proportional to:  $\exp\{-E_{\text{gap}}/kT\}$ .

In a semiconductor carrying current  $j$ , in a traverse magnetic field  $B$ , the **Hall effect** forces carriers of both polarities toward the same side of a crystal, creating an electric field  $E$ , with  $E = -jB/qN$ , where  $N$  is the dominant carrier density, and  $q$  is the carrier unit charge.

**Continuous wave functions in continuous space** are analyzed as wave functions in a crystal in the limit of zero atomic spacing. We replace summations with integrals, and define two infinite sets of basis states:  $|x\rangle$ , the state of a particle being at  $x$ ; and  $|p\rangle$ , the state of a particle having momentum  $p$ . In the  $|x\rangle$  basis representation, wave function  $\psi$  is:

$$\psi(x) = \langle x | \psi \rangle$$

Here  $\psi(x)$  is the *amplitude* for a particle to be at position  $x$  at time  $t$ , whereas  $|\psi\rangle$  is a particle *state* that might be spread across all  $x$ . Other restated equations are:

$$\begin{aligned} \langle \phi | \psi \rangle &= \int \langle \phi | x \rangle \langle x | \psi \rangle dx \\ \langle \phi | \psi \rangle &= \int \phi^*(x) \psi(x) dx \end{aligned}$$

The probability that an electron is between  $x$  and  $x+\Delta x$ , for small  $\Delta x$ , is:

$$\text{Prob}(x) \Delta x = |\psi(x)|^2 \Delta x$$

In the limit that  $\Delta x$  goes to zero,  $\text{Prob}(x)$  is the **probability density** of finding the particle near  $x$ .

Feynman chooses this normalization for momentum amplitudes:

$$\langle p|x\rangle = \langle x|p\rangle^* = \exp\{-ipx/\hbar\}$$
$$\text{Prob}(p) = |\langle p|\psi\rangle|^2 / h \text{ (not } \hbar)$$

The **Dirac delta function** is defined as the limit as  $\varepsilon$  goes to 0 of:

$$\delta(x) = \exp\{-x^2/\varepsilon^2\}/(\varepsilon\sqrt{\pi})$$
$$\int \delta(x) dx = 1, \text{ over all } x$$

**Schrödinger's equation** for non-relativistic velocities in energy potential  $V$  is:

$$1D: i\hbar d\psi/dt = -\hbar^2/2m \partial^2\psi/\partial x^2 + V\psi$$
$$3D: i\hbar d\psi/dt = -\hbar^2/2m \check{D}^2\psi + V\psi$$

Here  $\check{D} = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$  is the gradient  
and  $\check{D}^2 = \check{D} \cdot \check{D} = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$

## Symmetry & Conservation Laws

Feynman says: “*for each of the rules of symmetry there is a corresponding conservation law.*” This is called **Noether's Theorem**. Feynman says: “A physical system is *symmetric* with respect to the operation  $Q$  when  $Q$  commutes with  $U$  [or  $H$ ], the [operators for] the passage of time.” Mathematically, this means:  $QH=HQ$ . If a Hamiltonian  $H$  has a symmetry whose operator is  $Q$ , and  $Q$  has an eigenstate  $|\psi\rangle$ , then  $Q|\psi\rangle$  yields the same result forever, absent external interference. This establishes a conservation law:  $Q|\psi\rangle$  is conserved.

The following symmetry operations are universally valid.

- Translation in Space
- Translation in Time
- Spatial Rotation
- Motion at Constant Velocity
- Identical Particle Exchange
- Quantum Phase Change
- CPT Symmetry

The following symmetry operations are valid for all but the weak interaction.

- Spatial Reflection
- Time Reversal
- Electric Charge Polarity Reversal
- Matter & Antimatter Exchange

A symmetry operator operating on a non-degenerate eigenstate can only change its phase angle.

Symmetry under time translation implies **conservation of energy**.

Symmetry under x-axis translation implies **conservation of linear momentum** in the x-direction. The x-momentum operator is  $\hat{p}_x = \hbar k_x$ , with  $\partial\psi/\partial x = (p_x/\hbar)\psi$ . The same is true for y and z.

Symmetry under z-axis rotation implies **conservation of angular momentum** in the z-direction. Spin  $m$  in the z-direction is equivalent to angular momentum  $J_z = m\hbar$ . The same is true for x and y.

---

Let us first review what we learned about elementary particles in *Feynman Simplified 1A* Chapter 2, much of which was discovered after Feynman gave these lectures. The elementary particles, those believed not to be made of anything smaller, are listed here.

Figure 25-1  
Elementary Particles without Higgs

Each box in Figure 25-1 provides data for one particle, starting with the large one-letter symbol above the particle's name. The three numbers on the left side of each box are: the particle's mass, electric charge, and spin. Electric charges are stated in units of the proton charge. The small numbers in the figure may be hard to read, so we'll discuss them shortly.

The pleasingly symmetry of this 4-by-4 array is somewhat misleading. The right hand column contains four bosons: the photon, gluon, Z, and W. However, the boson on each row is not uniquely associated with the other particles on the same row. That's quite different from the other three columns, where all particles on the same row are intimately related.

The six quarks occupy the upper two rows of the left three columns. The top row contains the *up*, *charm*, and *top* quarks that all have electric charge  $+2/3$ . The second row contains the *down*, *strange*, and *bottom* quarks that all have charge  $-1/3$ . The quarks' names are entirely fanciful. In the subatomic realm, all spatial directions are equivalent; our customary notions of up, down, top, and bottom are meaningless. None of the quarks is more strange or less charming than the others.

The six leptons occupy the lower two rows of the left three columns. The bottom row contains the electron, muon, and tau that all have charge  $-1$ . The next row up contains the electron neutrino, muon neutrino, and tau neutrino that all have zero charge.

The four bosons have spin 1, while the quarks and leptons all have spin  $1/2$ . All quarks and leptons are fermions, the particles of matter. The four bosons are the exchange particles of forces: the photon for the electromagnetism; gluons for the strong force; and Z and W for the weak force.

Particle masses are measured in electron volts (eV), the energy an electron gains traversing a one-volt potential; MeV means million eV, and GeV means billion eV. In discussing particles, we employ the standard particle physics convention for the speed of light:  $c=1$ .

From left to right across the top row, the masses of u, c, t,  $\gamma$  are: 2.4 MeV, 1.27 GeV, 171.2 GeV, and zero. In the second row, the masses of d, s, b, g are: 4.8 MeV, 104 MeV, 4.2 GeV, and zero. In the third row, the neutrino masses are discussed below, and the Z mass is 91.2 GeV. In the fourth row, the masses of e,  $\mu$ ,  $\tau$ , and  $W^\pm$  are: 0.511 MeV, 105.7 MeV, 1.777 GeV, and 80.4 GeV.

Setting the neutrinos aside for the moment, the three generations of fermions are distinguished by their masses. In each row, the third generation particles are substantially more massive than the second generation, which in turn are substantially more massive than the first generation (left most column). Third generation particles can *decay* (transform) into second generation particles, which can decay into first generation particles. These decays happen rapidly; lifetimes range from millionths of a second to trillionths of a trillionth of a second. The first generation particles are stable (they appear to be eternal), because they have nothing lighter to decay into.

The masses of the three types of neutrinos remain unknown. We do know their masses are very small, and that the three mass values are all different. The values shown in Figure 25-1 are measured in particle reactions. Surprisingly, cosmological measurements provide a much more stringent limit: the sum of the three masses is less than 0.1 eV. In some sense, it is easier to measure the mass of the universe and subtract everything else than to measure the neutrino masses directly.

Not shown in Figure 25-1 is the Higgs boson, discovered in 2012. It should go with the other four bosons, but that would spoil the nice 4-by-4 array. The Higgs boson was widely hyped as the "God particle", but it is no more or less divine than any other particle. The Higgs boson is credited with being the source of the other elementary particles' masses. Particles that interact strongly with Higgs

become very massive; those interacting weakly with Higgs have low masses. Photons and gluons are massless because they don't interact at all with Higgs. The Higgs has zero electric charge, mass 125.2 GeV, and a lifetime of  $1.56 \times 10^{-22}$  sec. Its hallmark characteristic is that it has spin 0; this is unique among elementary particles, although many composite particles also have spin 0.

While elementary particles are thought to get their mass from the Higgs boson, the mass of composite particles comes from a quite different source. The proton mass, for example, is the sum of the masses of its constituent quarks plus their kinetic and interaction energies. Only the quark masses are due to Higgs effects; these amount to about 1% of the proton mass. Thus, about 99% of the mass of all normal matter is not due to Higgs effects, making it seem a bit less Godly.

Although not elementary, the proton and neutron are worth mentioning. A proton is comprised of two up quarks and one down quark; it has charge +1, mass 938.3 MeV, spin 1/2, and is stable (its lifetime is measured to be at least  $1.9 \times 10^{34}$  years). A neutron is comprised of two down quarks and one up quark; it has zero charge, mass 939.6 MeV, spin 1/2, and its lifetime is 881.5 seconds.

## **Lambda-Zero Decay**

Our first topic is the decay of the  $\Lambda^0$  ("lambda-zero"), and how angular momentum conservation affects that decay, which Feynman discusses in V3p17-11.

The  $\Lambda^0$  is a heavier cousin of the neutron. Its mass is 1115.7 MeV, about 19% more than a neutron. It has zero electric charge, spin 1/2, and even parity. The  $\Lambda^0$  decays through the weak force with a lifetime of  $2.63 \times 10^{-10}$  sec. This is a remarkably long lifetime compared with typical strong-force reaction times of  $10^{-24}$  sec.

The  $\Lambda^0$  is comprised of three quarks: one up, one down, and one strange. Its decay occurs in three stages, and involves a virtual  $W^-$  boson and an anti-up quark ( $\bar{u}$ ).

1. s decays to  $u + W^-$
2.  $W^-$  decays to  $d + \bar{u}$
3. u, d, u, d and  $\bar{u}$  quarks combine

In stage (3), the five quarks can combine in two different ways: 36% of the time they make a neutron (udd) and a neutral pion ( $u\bar{u}$ ); and 64% of the time they make a proton (uud) and a negative pion ( $d\bar{u}$ ), as shown in the two Feynman diagrams in Figure 25-2.

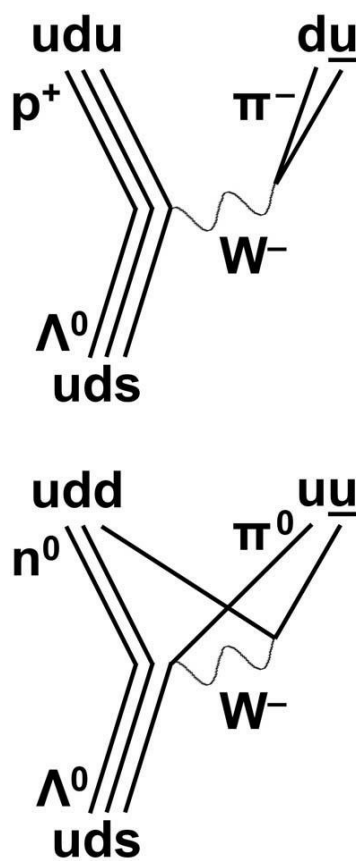


Figure 25-2 Two  $\Lambda^0$  Decay Modes

Let's consider the dominant decay mode that results in two charged particles:

$\Lambda^0$  decays to  $p^+$  and  $\pi^-$

Whenever one particle decays into two other particles, the decay products must have specific energy values in the center of mass (CM) reference frame, due to energy and momentum conservation. Let's see why. The total energy in the CM frame before and after the decay must be equal, hence:

$$m_{\Lambda} = m_p \gamma_p + m_{\pi} \gamma_{\pi}$$

Here,  $c=1$ ,  $\gamma=1/\sqrt{1-v^2}$  is the usual relativistic factor,  $v$  is velocity,  $m$  is mass, and the subscripts indicate to which particle each quantity pertains. In the CM frame the  $\Lambda^0$  is stationary, hence it has  $v=0$  and  $\gamma=1$ . The total momentum in the CM frame before and after the decay must also be equal, hence:

$$0 = m_p \gamma_p \mathbf{v}_p + m_{\pi} \gamma_{\pi} \mathbf{v}_{\pi}$$

Unfortunately, these equations don't reduce to anything simple, in general. In the  $\Lambda^0$  case, however,  $T_{\text{CM}}$ , the CM kinetic energy of the decay products, is quite low, as shown here:

$$\begin{aligned} m_{\Lambda} &= 1115.68 \text{ MeV} \\ m_p &= 938.27 \text{ MeV} \\ m_{\pi} &= 139.57 \text{ MeV} \end{aligned}$$

$$T_{\text{CM}} = 37.84 \text{ MeV}$$

Here the CM momenta of the proton and pion are each 100.58 MeV, and their velocities,  $0.107c$  and  $0.585c$  respectively, are only modestly relativistic.

The  $\Lambda^0$  and the proton have spin  $1/2$ , and the pion has spin  $0$ . Figure 25-3 shows this decay in the CM frame, with the z-axis set parallel to the spin of the lambda.

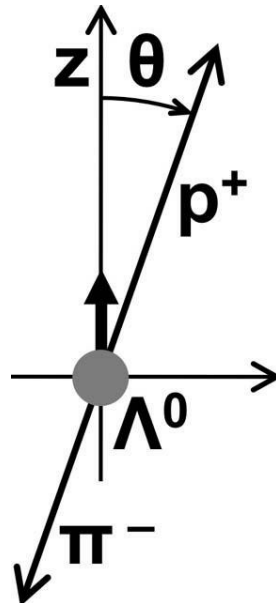


Figure 25-3 Lambda Decay Geometry

In Figure 25-3, the  $\Lambda^0$  spin is represented by the short vertical arrow, and  $\theta$  is the angle of the proton's velocity relative to the lambda spin direction, which is the z-axis.

From the conservation of momentum, in the CM frame, the pion's velocity must be exactly antiparallel to the proton's velocity. What we discover about the angle of the proton's velocity will immediately determine what the pion does as well. Let's find out what we can learn about angle  $\theta$ .

Consider the simplest case first (always a good starting point) of  $\theta=0$ , as shown in Figure 25-4.

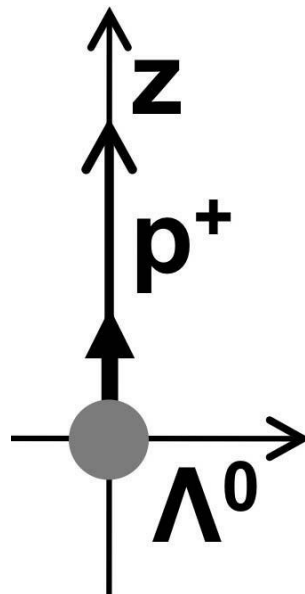


Figure 25-4 Lambda Decay Spin Up

When  $\theta=0$ , the proton is moving up the z-axis, and the pion is moving down the z-axis. Because both are moving along the z-axis,  $J_z$ , the z-component of angular momentum of both particles must be zero.

(Recall that  $J_z = xp_y - yp_x$ , and  $x=y=0$  along the  $z$ -axis.) Before the decay, the only angular momentum is the lambda spin, which equals  $\hbar/2$  in the  $+z$ -direction. After the decay, the only possible  $z$ -component of angular momentum,  $J_z$ , is the proton's spin, since the pion has zero spin. Therefore, to conserve  $J_z$ , the proton's spin must be up in the  $+z$ -direction; it cannot be spin down.

While the above logic is entirely valid, it may seem overly classical. Feynman therefore offers a second argument that is more quantum mechanical.

Define  $|\Lambda\rangle$  to be the state of the  $\Lambda^0$  before decaying, and  $|p@{\theta=0}\rangle$  to be the state of a proton moving at angle  $\theta$  in the CM frame. We need not be explicit about the pion, since its state is completely determined by the proton's state.

All known laws of nature, and therefore all the Hamiltonians we use to represent natural phenomena, are symmetric under rotation. Additionally,  $|\Lambda\rangle$  and  $|p@{\theta=0}\rangle$  are symmetric under rotations about the  $z$ -axis.

As defined in Chapter 23,  $R_z(\varnothing)$  is the operator for rotating the  $z$ -axis by angle  $\varnothing$ . According to what we discovered about symmetries in that chapter,  $R_z(\varnothing)$ ,  $|\Lambda\rangle$ , and  $|p@{\theta=0}\rangle$  satisfy these equations:

$$\begin{aligned} R_z(\varnothing) |\Lambda\rangle &= \exp\{i\varnothing/2\} |\Lambda\rangle \\ R_z(\varnothing) |p@{\theta=0}\rangle &= \exp\{i\varnothing/2\} |p@{\theta=0}\rangle \end{aligned}$$

The key points demonstrated by these equations are: (1)  $R_z(\varnothing)$  can change an eigenstate by only a phase angle; (2) this phase angle is a constant of the motion that never changes regardless of how the state evolves, since  $R_z H = H R_z$ ; and (3) the phase angle for spin  $1/2$  particles is  $\varnothing/2$ , as given in the rotation tables at the end of Chapter 17.

We also discovered in Chapter 23 that the phase angle  $\varnothing/2$  is proportional to  $J_z$ . These equations prove, therefore, that angular momentum is conserved and that the proton must be spin up if  $\theta=0$ .

Define  $A$  to be the amplitude for the lambda to decay to a proton at  $\theta=0$  with spin in the  $+z$ -direction, that is:

$$A = \langle p@{\theta=0}, s=+z | H | \Lambda, s=+z \rangle$$

Here we extended the state definitions to include spin orientation  $s$ .

Now consider the same decay, but this time with the lambda spin in the  $-z$ -direction, as shown in Figure 25-5.

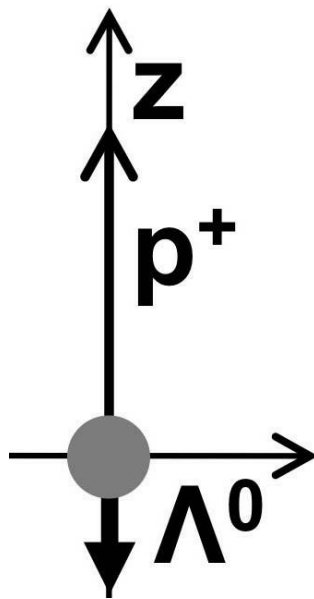


Figure 25-5 Lambda Decay Spin Down

The same logic, with the opposite sign, shows that for  $\theta=0$  the proton must be spin down to conserve angular momentum. Define B to be the amplitude of this decay:

$$B = \langle p @ \theta=0, s=-z | H | \Lambda, s=-z \rangle$$

If the decay process were symmetric under reflection — if it conserved parity — that would ensure  $B=\pm A$ , as we now demonstrate. If the configuration of Figure 25-4 were reflected in the horizontal plane, the z-axis and the proton direction would flip, pointing downward, parallel to the lambda spin, resulting in the same configuration as in Figure 25-4, whose amplitude is A. Since the parity operator can only change an eigenstate by a factor of  $\pm 1$ , this would mean  $B=\pm A$ . Recall that axial vectors, like spin, do not change when reflected in a plane perpendicular to their direction (see *Feynman Simplified 1D* Chapter 49).

However, since this decay process requires a weak interaction, reflection symmetry and parity conservation are not guaranteed: A and B have no definite relationship.

To review:  $\theta$  is the angle between the  $\Lambda^0$  spin and the proton's velocity. We define A and B to be the amplitudes for this decay at these special values of  $\theta$ :

A = amplitude at  $\theta = 0$

B = amplitude at  $\theta = \pi$

At both angles, the proton and  $\Lambda^0$  spins must be parallel to conserve angular momentum. Our next step is to calculate the amplitudes for intermediate angles. As Feynman says, in V3p17-13, with: “these two amplitudes we can find out all we want to know about the angular distribution of the disintegration.”

We achieve this by defining a new axis called  $z^*$  that is rotated by angle  $\theta$  from the z-axis about the axis perpendicular to the screen (call that the y-axis). This rotation is performed by the  $R_y(\theta)$  operator, defined in Chapter 17. A  $\Lambda^0$  with spin up along the z-axis can either be spin up or spin down along the  $z^*$ -axis with amplitudes given by  $R_y(\theta)$ :

$$\langle \Lambda s=+z^* | \Lambda s=+z \rangle = +\cos(\theta/2)$$

$$\langle \Lambda s=-z^* | \Lambda s=+z \rangle = -\sin(\theta/2)$$

We next calculate the amplitude for the proton to be moving in the  $+z^*$ -direction for each spin possibility, as shown in Figure 25-6.

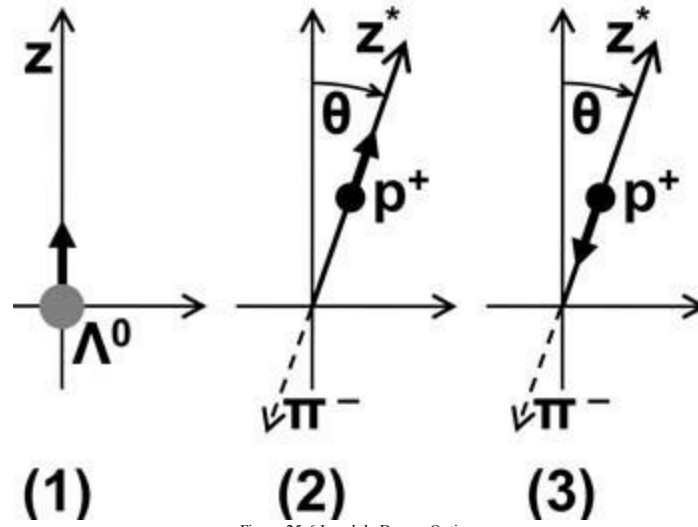


Figure 25-6 Lambda Decay Options

Part (1) of the figure shows the initial  $\Lambda^0$  spin orientation in the  $+z$ -direction. Part (2) and part (3) show the proton direction of motion in the  $+z^*$ -direction, and at angle  $\theta$  relative to the  $z$ -axis. In part (2) the proton spin is in the  $+z^*$ -direction, and in part (3) the proton spin is in the  $-z^*$ -direction.

We will calculate part (2) first. In principle this could occur in two indistinguishable ways, so we must sum the amplitudes for each way. The two ways correspond to  $\Lambda^0$  spin up or spin down in the  $+z^*$ -direction. Each of these two amplitudes is itself the product of two other amplitudes: the amplitude for  $\Lambda^0$  to have that spin; and the amplitude that a  $\Lambda^0$  of that spin would decay to a proton with spin in the  $+z^*$ -direction. We therefore get this quite long expression:

$$\begin{aligned} &\langle p \text{ along } z^*, s=+z^* | \Lambda s=+z \rangle \\ &= \langle p, s=+z^* | H | \Lambda s=+z^* \rangle \langle \Lambda s=+z^* | \Lambda s=+z \rangle \\ &\quad + \langle p, s=+z^* | H | \Lambda s=-z^* \rangle \langle \Lambda s=-z^* | \Lambda s=+z \rangle \end{aligned}$$

We know from above that the Hamiltonian component in the lowest line is zero, due to conservation of angular momentum: a proton emitted along the  $\Lambda^0$  spin axis must have spin parallel to the  $\Lambda^0$ . In the middle line, the Hamiltonian component equals  $A$  and the rotation factor equals  $\cos(\theta/2)$ . This reduces the long expression to:

$$\langle p \text{ along } z^*, s=+z^* | \Lambda s=+z \rangle = +A \cos(\theta/2)$$

Similarly, for part (3), the proton moving in the  $+z^*$ -direction with spin in the  $-z^*$ -direction:

$$\begin{aligned} &\langle p \text{ along } z^*, s=-z^* | \Lambda s=+z \rangle \\ &= \langle p, s=-z^* | H | \Lambda s=+z^* \rangle \langle \Lambda s=+z^* | \Lambda s=+z \rangle \\ &\quad + \langle p, s=-z^* | H | \Lambda s=-z^* \rangle \langle \Lambda s=-z^* | \Lambda s=+z \rangle \end{aligned}$$

Here the Hamiltonian in the middle line is zero by angular momentum conservation, reducing the expression to:

$$\langle p \text{ along } z^*, s=-z^* | \Lambda s=+z \rangle = -B \sin(\theta/2)$$

These protons are moving in the  $+z^*$ -direction, which is at angle  $\theta$  relative to the  $z$ -axis.

Now consider an experiment that detects protons from  $\Lambda^0$  decay and measures the angle between the proton velocities and the  $\Lambda^0$  spin in the CM frame. Assume the experiment does not measure the direction of the proton spin: it tallies all protons regardless of whether they are spin up or spin down.

Here is your pop quiz: is the probability of a proton at angle  $\theta$  equal to the square of the sum of amplitudes, or the sum of the squares of amplitudes?

dee-dee, dee-dee, dee-dee, dee-dee...ready to check your answer?

Protons with opposite spins are not identical; they are distinguishable even if we do not bother to distinguish them. We therefore sum squares.

$$\text{Prob}(p@ \theta) = |A|^2 \cos^2(\theta/2) + |B|^2 \sin^2(\theta/2)$$

Using the half-angle formulae, we can rewrite this as:

$$\begin{aligned} \text{Prob}(p@ \theta) &= \beta (1 + \alpha \cos \theta) \\ \alpha &= (|A|^2 - |B|^2) / (|A|^2 + |B|^2) \\ \beta &= (|A|^2 + |B|^2) / 2 \end{aligned}$$

In V3p17-14, Feynman says experiments confirm this cosine dependence, and provide a measured value of  $\alpha$  that is  $-0.62 \pm 0.05$ . From that we can find the  $|A/B|$  ratio.

$$\begin{aligned} \alpha &= (|A/B|^2 - 1) / (|A/B|^2 + 1) \\ \alpha + \alpha |A/B|^2 &= |A/B|^2 - 1 \\ \alpha + 1 &= |A/B|^2 (1 - \alpha) \\ |A/B|^2 &= (\alpha + 1) / (1 - \alpha) \\ |A/B|^2 &= (0.38 \pm 0.05) / (1.62 - [\pm 0.05]) \\ |A/B|^2 &= 0.235 \pm 0.038 \\ |A/B| &= 0.485 \pm 0.040 \end{aligned}$$

Therefore  $B$  is about twice as large as  $A$ , confirming that parity symmetry is violated in the  $\Lambda^0$  decay.

Feynman then seeks the probability that the proton will be spin up along the original  $z$ -axis. We already have the essential pieces that answer this question. We have the amplitudes for the proton to be spin up and spin down along the  $z^*$ -axis. Applying  $R_y(-\theta)$  projects these amplitudes back onto the  $z$ -axis.

$$\begin{aligned} |p, s=+z \rangle &= \langle +z | R_y(-\theta) | +z^* \rangle |p, s=+z^* \rangle \\ &+ \langle +z | R_y(-\theta) | -z^* \rangle |p, s=-z^* \rangle \end{aligned}$$

$$|p, s=+z\rangle = [\cos(\theta/2)] [A\cos(\theta/2)] \\ + [-\sin(\theta/2)] [-B\sin(\theta/2)]$$

$$|p, s=+z\rangle = A \cos^2(\theta/2) + B \sin^2(\theta/2)$$

The probability is the magnitude of this amplitude squared.

## Chapter 25 Review: Key Ideas

The  $\Lambda^0$  (“lambda-zero”) is a more massive cousin of the neutron in which one of the neutron’s down quarks is replaced by a strange quark. The  $\Lambda^0$  has a remarkably long lifetime of  $2.63 \times 10^{-10}$  seconds. It decays through the weak force, primarily to a proton and a negative pion.

In the CM frame of that decay, let  $\theta$  be the angle between the proton’s velocity and the  $\Lambda^0$  spin. Let  $A$  be that amplitude at  $\theta=0$ , and  $B$  be that amplitude at  $\theta=\pi$ . At both of these angles, the proton and  $\Lambda^0$  spins must be parallel, due to the conservation of angular momentum. The probability of the proton’s velocity being at an intermediate angle  $\theta$  is:

$$\text{Prob}(p@ \theta) = |A|^2 \cos^2(\theta/2) + |B|^2 \sin^2(\theta/2)$$

Experiments determine that  $|A/B| = 0.485 \pm 0.40$ . Since  $A/B$  is not  $\pm 1$ , parity is not conserved in this weak interaction.

# Chapter 26

## Particle Physics: Neutral Kaons

---

The exotic properties of neutral kaons present some of the most beautiful demonstrations of quantum principles in particle physics. Their once inexplicable properties drove physicists to embrace “strangeness.”

In V3p11-12, Feynman begins a long discussion of how little was known about neutral kaons 50 years ago. Rather than detailing everything we didn’t know then, let’s instead discuss all that we do know now.

### Particle Conservation Laws

Recall our discussion of elementary particles in the prior chapter. These are the particles that we believe are the most fundamental pieces of nature, those that are not made of anything “smaller.” These include: four force-carrying bosons; the Higgs boson; twelve fermions; and the corresponding twelve antifermions. The fermions are further subdivided into six quarks that participate in the strong force, and six leptons that do not.

Historically, physicists identified a steadily increasing number of rules governing the creation and annihilation of various types of elementary particles in different types of reactions. For example, we have long believed that the total number of leptons is always conserved. *Lepton conservation* means no reaction ever changes the total number of leptons of all six types minus the total number of antileptons of all six types. No evidence has ever been found that invalidates the absolute conservation of leptons.

With our modern understanding of particle physics, called the Standard Model of Particle Physics, we can state these conservation laws more succinctly:

In all processes driven by either the strong force or the electromagnetic force, the total number of fermions minus antifermions of each type is always conserved. The number of bosons is not conserved, neither within one type nor summed over all types.

This means the total number of up quarks minus anti-up quarks never changes, and similarly for each of the other eleven types of elementary fermions. This amounts to twelve separate conservation laws. In all these particle conservation laws, one sums the number of particles and subtracts the number of corresponding antiparticles.

In the past, a major principle of particle physics was the conservation of baryons (protons, neutrons, and their heavier cousins). This conservation law explained why the following reaction has never been observed:

$$p^+ + p^+ \text{ never goes to } \pi^+ + \pi^+ + \pi^0$$

Above, we have 2 baryons on the left and none on the right; hence this reaction violates the conservation of baryon number.

While baryon number conservation is still a valid principle, it isn't the most insightful description of nature. Baryon conservation is simply a consequence of the conservation of quarks of each type, and the way that quarks combine.

A  $\pi^+$  has 1 up quark and 1 anti-down quark, while a  $\pi^0$  is a linear combination of up+anti-up and down+anti-down. The above reaction has 4 up quarks on the left and only 2 on the right. It also has 2 down quarks on the left and  $-2$  down quarks on the right. The reaction is doubly forbidden because it violates the conservation of two types of quarks.

If that were the end of the story life would be simple, too simple to allow the existence of interesting things such as stars, planets, and life.

The complication, and our opportunity for existence, comes from the weak force. Only the weak force is known to change one type of elementary fermion into another. This enables nuclear fusion that powers the stars and produces all elements heavier than helium, which is quite important for carbon-based life forms. Since changing fermion types can occur only through the weak force, such reactions proceed at vastly lower rates (less probability per second) than reactions driven by the strong or electromagnetic forces.

The ability of the weak force to change one fermion into another is not universal. Weak force phenomena are typically a bit complicated.

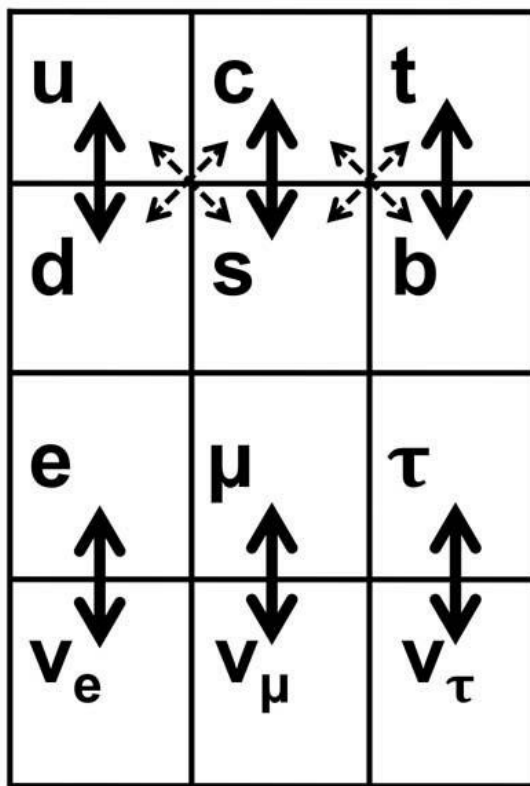


Figure 26-1. Weak & Weaker Transitions

In Figure 26-1, the bold arrows indicate transitions between fermion types that the weak force enables at relatively high rates — high relative to other weak interactions. These include transitions between up and down quarks, and between electrons and electron-neutrinos.

The dotted arrows indicate transitions between fermion types that the weak force enables, but at much lower rates. These include transitions between up and strange quarks, and between down and charm quarks.

Note that none of these transitions changes the total number of quarks minus antiquarks, nor the total number of leptons minus antileptons. In appropriate circumstances, all these transitions can proceed in either direction.

None of these transitions can occur in isolation. When a down quark transitions to an up quark, electric charge increases by +1. That transition can only occur in tandem with another reaction that changes electric charge by −1, such as an electron-neutrino transitioning to an electron. Additionally, energy must be conserved in all reactions. Since a charm quark has more mass than a down quark, a down-to-charm transition must occur in tandem with another reaction that supplies the required energy.

We also know that transitions occur in which neutrinos change types. Those reactions are not well-understood at this time, but they most probably involve the weak force.

Particle conservation laws require that the number of strange quarks is conserved in strong or electromagnetic force reactions. This is why the  $\Lambda^0$  decay discussed in the prior chapter proceeds through the weak force. The  $\Lambda^0$  is the least massive three-quark particle that contains a strange quark. In every particle decay without exception, the sum of the masses of the decay products must be less than the mass of the decaying particle. (If that isn't clear, think about the total energy in the original

particle's center of mass frame before and after the decay.) This means the  $\Lambda^0$  can only decay to particles that have no strange quarks. Because the number of strange quarks has to change in  $\Lambda^0$  decay, this decay can only occur through the weak force.

Before physicists knew about quarks, Murray Gell-Mann and Abraham Pais explained the long  $\Lambda^0$  lifetime by introducing a concept called *strangeness*. Each particle was assigned a *strangeness quantum number*  $S$ , and a new law required the conservation of strangeness in all but weak force reactions. The  $\Lambda^0$  has strangeness  $-1$ , protons and pions have strangeness zero, so the  $\Lambda^0$  decay results in a strangeness change of  $\Delta S = +1$ , thus requiring the weak force and greatly extending the  $\Lambda^0$  lifetime.

The strangeness concept succeeded in explaining why some particles had such long lifetimes. It also explained why certain reactions occurred and other seemingly similar reactions did not. This was an important step toward the later development of the quark model, the concept that all strongly interacting particles are composed of quarks.

Unfortunately, the sign convention that Gell-Mann and Pais selected resulted in assigning the strange quark  $S = -1$ . This is much like the unfortunate sign convention for electric charge in which positive current flowing left is actually due to negative electrons flowing right. In both cases, by the time physicists developed complete theories, these sign conventions were too deeply ingrained to be reversed.

When additional types of quarks were discovered, the logical extension of strangeness resulted in quantum numbers and conservation laws for *charmness*, *topness*, and *bottomness* — a linguist and political-correctness nightmare.

It is far better to say that each fermion type is conserved except in weak interactions that conserve only total quark number and total lepton number. This simpler rule eliminates the need for the separate conservation laws for strangeness, charmness, topness, and bottomness. I often refer to strangeness conservation for clarity, for historical reasons, and because that is common usage.

## Kaons & The Strong Force

Kaons, once called K-mesons, are copiously produced in strong force reactions. The four types of kaons produced in strong force reactions have the following symbols, quark contents, strangeness quantum numbers, and masses:

$K^+$ :  $u\bar{s}$ ;  $S = +1$ ; 493.67 MeV

$K^-$ :  $s\bar{u}$ ;  $S = -1$ ; 493.67 MeV

$K^0$ :  $d\bar{s}$ ;  $S = +1$ ; 497.65 MeV

$\bar{K}^0$ :  $s\bar{d}$ ;  $S = -1$ ; 497.65 MeV

The superscripts indicate the particles' electric charge. Note that the  $K^+$  and  $K^-$  are antiparticles of one another, as evidenced by their opposite quark compositions. Similarly, the  $K^0$  and  $\bar{K}^0$  ("K-zero-bar") are different particles that are antiparticles of one another, because they contain opposite quarks. (In physics texts and research papers, antiquarks and neutral antiparticles are denoted with

bars above their symbols, but that isn't possible in an eBook, so I use underlines.)

Since each kaon contains one  $s$  or  $\bar{s}$  quark, and since strong force reactions conserve the number of strange quarks, strong reactions must produce kaons in conjunction with other particles containing the opposite  $\bar{s}$  or  $s$  quarks. Two such reactions are:

$$\pi^- + p^+ \text{ goes to } \Lambda^0 + K^0$$

$$n^0 + p^+ \text{ goes to } p^+ + p^+ + K^- + K^0$$

In the first reaction,  $S=0+0$  on the left and  $-1+1$  on the right, conserving strangeness. In the second reaction,  $S=0+0$  on the left and  $0+0-1+1$  on the right, again conserving strangeness.

Strangeness conservation in strong interactions also explains why:

$$\underline{K}^0 + p^+ \text{ goes to } \Lambda^0 + \pi^+$$

$$K^0 + p^+ \text{ never goes to } \Lambda^0 + \pi^+$$

In the first reaction,  $S=-1+0$  on the left and  $-1+0$  on the right. In the second reaction,  $S=+1+0$  on the left and  $-1+0$  on the right, which is forbidden in a strong force reaction.

## Kaon Decays

By the norms of particle physics, the  $K^+$  and  $K^-$  decay very slowly, with a lifetime of  $1.238 \times 10^{-8}$  seconds. Their two principle decay modes are:

$$K^\pm \longrightarrow \mu^\pm + \text{neutrino}$$

$$K^\pm \longrightarrow \pi^\pm + \pi^0$$

The long lifetime confirms that these are weak reactions in which strangeness is not conserved.

The neutral kaons are far more interesting. They can decay via the weak force into:

$$K^0 \text{ or } \underline{K}^0 \longrightarrow \pi^+ + \pi^- \text{ or } \pi^0 + \pi^0$$

As Gell-Mann and Pais pointed out, since both neutral kaons can decay to a pair of pions, there is a non-zero amplitude for each of the particles to transform into the other, with the weak force acting twice, through these reactions:

$$K^0 \longrightarrow 2 \text{ pions} \longrightarrow \underline{K}^0$$

$$\underline{K}^0 \longrightarrow 2 \text{ pions} \longrightarrow K^0$$

In my copy of the *Feynman Lectures*, which is the book's first edition, the equations in V3p11-17 contain a number of errors. In a recent edition, the derivation is done substantially differently, but in

a manner that raises technical issues that I discuss in the last section of this chapter. And, after 50 years, all editions of these lectures are considerably outdated.

My approach will therefore deviate from the *Feynman Lectures*, and will instead reflect our modern understanding of neutral kaons.

Gell-Mann and Pais realized that  $K^0$  and  $\bar{K}^0$  are not two independent particles. Instead, their decay processes unite them as the basis states of one combined two-state system. The neutral kaons were the first discovered example of a more general phenomenon: *neutral particle oscillations*, which have also been observed with B's, D's, and neutrinos.

In all the two-state systems we have examined before, the components of the Hamiltonian were real numbers. This is because those systems did not decay; the total number of particles in those systems never changed. Now we have something different: neutral kaons do decay. Therefore the solutions to the kaon two-state system must decrease exponentially over time. We can accommodate decay by including imaginary amplitudes in the Hamiltonian.

Let's examine that mathematically. For a stable system, the diagonal components of the Hamiltonian are real numbers that equal the energy of the basis states, thus:

$$\begin{aligned} H_{nn} &= E_0 \\ i\hbar \frac{dC_n}{dt} &= E_0 C_n \\ C_n &= \exp\{E_0 t / i\hbar\} \end{aligned}$$

This solution rotates within the complex plane, with a changing phase angle and a constant magnitude.

Now consider a complex diagonal component of the Hamiltonian:

$$\begin{aligned} H_{nn} &= E_0 - i\beta \\ i\hbar \frac{dC_n}{dt} &= (E_0 - i\beta) C_n \\ C_n &= \exp\{E_0 t / i\hbar\} \exp\{-\beta t / \hbar\} \end{aligned}$$

We see that  $E_0$  leads to an oscillating factor, the basis states “sloshing back and forth”, while  $-i\beta$  provides the decaying exponential factor. This is why we now introduce complex amplitudes.

We begin analyzing the neutral kaon system by selecting basis states  $|K^0\rangle$  and  $|\bar{K}^0\rangle$ , and defining:

$$\begin{aligned} C_+ &= \langle K^0 | \psi \rangle \\ C_- &= \langle \bar{K}^0 | \psi \rangle \end{aligned}$$

The plus and minus signs match the strangeness quantum numbers of the  $K^0$  and  $\bar{K}^0$ .

The Hamiltonian, with complex amplitudes, is:

$$\begin{aligned} H_{++} &= E_0 - i\beta \\ H_{+-} &= a - i\alpha \end{aligned}$$

$$H_{-+} = a - i\alpha$$

$$H_{--} = E_0 - i\beta$$

where:

$$a - i\alpha = \langle K^0 | \underline{K}^0 \rangle = \langle \underline{K}^0 | K^0 \rangle$$

Following our well-trod path for analyzing two-state systems, the Hamiltonian equations are:

$$i\hbar dC_+/dt = (E_0 - i\beta) C_+ + (a - i\alpha) C_-$$

$$i\hbar dC_-/dt = (E_0 - i\beta) C_- + (a - i\alpha) C_+$$

As before, adding and subtracting the above equations yields the stationary solutions.

$$i\hbar d(C_+ + C_-)/dt = (E_0 - i\beta)(C_+ + C_-) + (a - i\alpha)(C_+ + C_-)$$

$$i\hbar d(C_+ - C_-)/dt = (E_0 - i\beta)(C_+ - C_-) - (a - i\alpha)(C_+ - C_-)$$

For the stationary states, define:

$$|K_1\rangle = (|K^0\rangle - |\underline{K}^0\rangle) / \sqrt{2}$$

$$|K_2\rangle = (|K^0\rangle + |\underline{K}^0\rangle) / \sqrt{2}$$

$$C_1 = (C_+ - C_-) / \sqrt{2}$$

$$C_2 = (C_+ + C_-) / \sqrt{2}$$

The energy levels of the “stationary” states are:

$$E_1 = E_0 - a$$

$$E_2 = E_0 + a$$

With decaying particles, no state is completely “stationary.” What we are referring to are those states whose basis state amplitudes maintain constant ratios; the states that maintain their essential character as their overall magnitudes decrease exponentially.

In the kaon rest frame, these energy levels must match the measured masses, which yields:

$$E_0 = 497.65 \text{ MeV}$$

$$2a = 3.5 \times 10^{-12} \text{ MeV}$$

The  $K_2$  is minutely more massive than an average  $K_1$ . We must say “average  $K_1$ ” because unstable particles do not have definite masses. The Uncertainty Principle states that the precision of an energy measurement  $\Delta E$ , measured over a time interval  $\Delta t$ , can never be better than:  $\Delta E \Delta t = \hbar/2$ . The time interval for a mass measurement of a  $K_1$  is limited to its lifetime. Hence the least mass uncertainty occurs when the  $K_1$  mass has a Gaussian distribution whose full width is  $\Delta m_1 = \hbar/\Delta t$ . With the measured  $K_1$  lifetime quoted below,  $\Delta m_1 = 7.65 \times 10^{-12} \text{ MeV}$ . This means the variation in mass of a  $K_1$  is about twice the  $K_2$ – $K_1$  average mass difference  $\Delta m_{12}$ . This is illustrated in Figure 26-2.

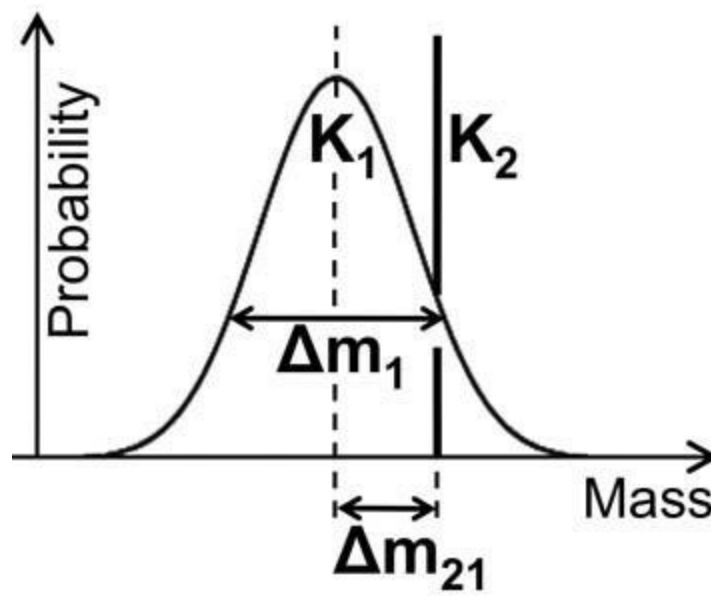


Figure 26-2 K<sub>1</sub>, K<sub>2</sub> Mass Distributions

The K<sub>2</sub> mass distribution is represented in the figure by a bold vertical line, because its full width of  $1.3 \times 10^{-14}$  MeV is about 600 times smaller than the K<sub>1</sub> mass distribution.

The stationary states have the following time dependence:

$$C_1(t) = C_1(0) \exp\{E_1 t / i\hbar\} \exp\{-(\beta - \alpha)t / \hbar\}$$

$$C_2(t) = C_2(0) \exp\{E_2 t / i\hbar\} \exp\{-(\beta + \alpha)t / \hbar\}$$

We can determine  $\beta$  and  $\alpha$  from the measured mean lifetimes.

$$K_1: t_1 = 8.954 \times 10^{-11} \text{ sec}$$

$$K_2: t_2 = 5.116 \times 10^{-8} \text{ sec}$$

$$(\beta - \alpha) / \hbar = 1 / 8.954 \times 10^{-11} \text{ sec} = 1.117 \times 10^{+10} / \text{sec}$$

$$(\beta + \alpha) / \hbar = 1 / 5.116 \times 10^{-8} \text{ sec} = 1.955 \times 10^{+7} / \text{sec}$$

$$\beta / \hbar = 5.594 \times 10^{+9} / \text{sec}$$

$$\alpha / \hbar = 5.574 \times 10^{+9} / \text{sec}$$

We can also express the tiny mass difference in the same units.

$$2a / \hbar = 3.5 \times 10^{-12} \text{ MeV} / 6.582 \times 10^{-22} \text{ MeV-sec}$$

$$2a / \hbar = 5.32 \times 10^{+9} / \text{sec}$$

$$t_{\Delta m} = \hbar / 2a = 1.88 \times 10^{-10} \text{ sec}$$

There are three time constants in this system: (1) the K<sub>1</sub> lifetime  $t_1$ ; (2) the K<sub>2</sub> lifetime  $t_2$ ; and (3) the two-basis-state oscillation time  $t_{\Delta m}$ . The measured time constants have these relationships:

$$t_{\Delta m} = 2.1 t_1$$

$$t_2 = 571 t_1$$

We can rewrite the stationary states highlighting these time constants.

$$C_1(t) = C_1(0) \exp\{-imt/\hbar\} \exp\{-it/t_{\Delta m}\} \exp\{-t/t_1\}$$

$$C_2(t) = C_2(0) \exp\{-imt/\hbar\} \exp\{+it/t_{\Delta m}\} \exp\{-t/t_2\}$$

Since the decay time constant  $t_1$  is less than the oscillation time constant  $t_{\Delta m}$ , this system is an over-damped harmonic oscillator, as described in *Feynman Simplified 1B* Chapter 14.  $C_1$  rapidly vanishes, leaving only  $C_2$ .

From the stationary state solutions, we can calculate the basis state amplitudes.

$$C_1 = (C_+ - C_-) / \sqrt{2}$$

$$C_2 = (C_+ + C_-) / \sqrt{2}$$

$$C_+ = (C_1 + C_2) / \sqrt{2}$$

$$C_- = (C_2 - C_1) / \sqrt{2}$$

For  $t \gg t_1$ ,  $C_1$  vanishes and we are left with:

$$C_+ = +C_2 / \sqrt{2}$$

$$C_- = +C_2 / \sqrt{2}$$

A  $K^0$  produced in a strong reaction at  $t=0$  has:

$$C_+(t=0) = 1$$

$$C_-(t=0) = 0$$

$$C_1(0) = C_2(0) = 1/\sqrt{2}$$

At a time  $t$ , with  $t_1 \ll t \ll t_2$ , this particle becomes:

$$C_+(t) = +\exp\{-imt/\hbar\} \exp\{-it/t_{\Delta m}\} / 2$$

$$C_-(t) = +\exp\{-imt/\hbar\} \exp\{-it/t_{\Delta m}\} / 2$$

What all this means is that the original  $K^0$  has a 50% probability of decaying into a pair of pions, and a 50% probability of becoming a symmetric superposition of  $K^0$  and  $\bar{K}^0$  — its original self and its antiparticle in equal measures: a quantum version of Dr. Jekyll and Mr. Hyde. A  $\bar{K}^0$  produced in a strong reaction has exactly the same fate.

The  $K_L$  slowly decays while remaining a superposition of matter and antimatter. Its quark composition is always  $(d\bar{s}+s\bar{d})/\sqrt{2}$ .

Recall that CP symmetry is the combination of charge polarity reversal (C) plus spatial inversion (P), which amounts to the exchange of matter and antimatter. Like parity P, elementary particles (generally) are eigenstates of definite CP — when operated on by CP, their amplitudes are multiplied by either +1 or -1. This means particles have a conserved CP quantum number that is either +1 or -1. If CP is a universal symmetry of nature, the CP quantum number of all closed systems never changes.

Also like parity, the CP quantum number of two particles equals the product of the CP quantum

numbers of each particle separately. Pions have  $CP=-1$ , but a pair of pions has  $CP=(-1)^2 = +1$ . The  $CP$  of a  $K_1$  is  $+1$ , while the  $CP$  of a  $K_2$  is  $-1$ . This means a  $K_1$  can decay to two pions, but a  $K_2$  cannot, if  $CP$  is conserved.

When a  $K_1$  decays to two charged pions, about 218 MeV of mass energy is released. By comparison, when a  $K_2$  decays to three charged pions, which have  $CP=-1$ , only about 79 MeV of mass energy is released. This means the  $K_2$  decay proceeds at a much lower rate than the  $K_1$  decay.

Now we come to more recent discoveries.

## Kaon Regeneration

A remarkable quantum phenomenon occurs when a beam of long-lived neutral kaons passes through a material object.

In 1955, Abraham Pais and Oreste Piccioni predicted that a  $K_2$  beam traversing matter *regenerates* a coherent  $K_1$  beam. This is because normal matter preferentially absorbs the  $\underline{K}^0$  component of a  $K_2$  beam, as we will now explore.

From above, absent an absorber, the basis state composition of a  $K_2$  beam is:

$$|K_2\rangle = (|K^0\rangle + |\underline{K}^0\rangle) / \sqrt{2}$$

For simplicity, assume that in passing through a thin block of matter, a fraction  $f$  of the  $\underline{K}^0$  component of this beam is absorbed and none of its  $K^0$  component is absorbed. The basis state composition of the beam becomes (recalling that  $|K^0\rangle = (|K_1\rangle + |K_2\rangle) / \sqrt{2}$ ):

$$\begin{aligned} & (|K^0\rangle + (1-f)|\underline{K}^0\rangle) / \sqrt{2} \\ &= \{(1-f)(|K^0\rangle + |\underline{K}^0\rangle) / \sqrt{2}\} + \{f|K^0\rangle\} \\ &= \{(1-f)|K_2\rangle\} + \{f(|K_1\rangle + |K_2\rangle) / \sqrt{2}\} \\ &= (1-f+f/\sqrt{2})|K_2\rangle + f|K_1\rangle / \sqrt{2} \end{aligned}$$

The key feature is the reappearance, the *regeneration*, of  $K_1$  particles. Half of these regenerated  $K_1$ 's rapidly decay to pairs of pions, leaving the other half to rejoin the  $K_2$  beam.

In 1960, an experiment lead by Oreste Piccioni observed this kaon regeneration process.

An old physicists' adage laments: "Yesterday's *sensation* is today's *calibration* and tomorrow's *contamination*." As science continually advances, our machines become more powerful. Phenomena that were once barely detectable eventually are seen routinely, and become useful benchmarks. Even later, they become so ubiquitous that they must be methodically eliminated to unmask even rarer phenomena. In brief: even thrilling discoveries become boring and burdensome.

As an example, less than a decade after my father first observed kaon regeneration, the pion-muon discriminator in our  $CP$ -violation experiment was calibrated using regenerated kaons. Once

calibrated, we took special measures to ensure unintended kaon regeneration did not contaminate our data.

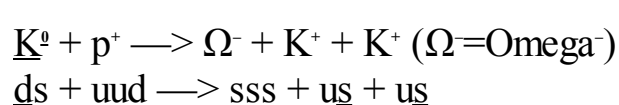
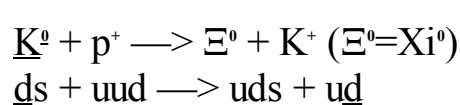
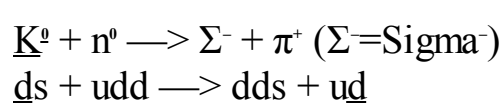
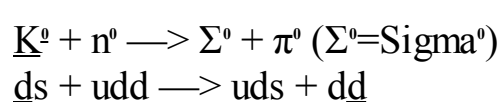
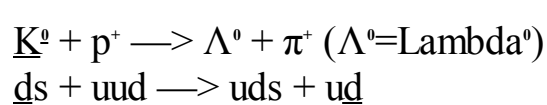
Now, let's discuss why matter preferentially absorbs  $\underline{K}^0$  and not  $K^0$ . Neutral kaons are absorbed by matter through their strong force reactions with the protons and neutrons that all normal matter contains. Recall that the strong force conserves the net number of each type of quark, and that the  $\underline{K}^0$  contains one strange quark, while the  $K^0$  contains one antistrange quark. By "net" number, we mean the number of particles minus the number of antiparticles.

Protons and neutrons are the two least massive baryons, a class of particles that have a net quark number equal to +3. Protons and neutrons contain up and down quarks, but no strange quarks.

Some of the more massive baryons, called *hyperons*, contain one or more strange quarks, but none contain any antistrange quarks, which count as  $-1$  quark. A baryon with one antistrange quark would need four normal quarks to reach the required net quark number of +3. While a *pentaquark*, four quarks plus one antiquark, is not impossible, none has ever been observed. Also, a pentaquark would necessarily contain at least one quark-antiquark pair, making it extremely short-lived.

All this means that many types of hyperons can be produced by  $\underline{K}^0$  collisions with matter, but cannot be produced by similar  $K^0$  collisions. With more reaction opportunities,  $\underline{K}^0$  has a much greater absorption rate in normal matter than  $K^0$ .

Examples of hyperon production in  $\underline{K}^0$  collisions are listed below, with each pair of lines describing one reaction. In each pair, the first line shows the reaction symbolically (with the hyperon name), while the second line shows the quark content of each particle.



## CP Violation

In 1964, James Cronin and Val Fitch observed the long-lived neutral kaon decaying to two charged pions, in violation of CP conservation. For this discovery, they were awarded the 1980 Nobel Prize.

My own thesis experiment later discovered a CP-violating asymmetry in long-lived neutral kaon decays to:

$$\pi^+ \mu^- \underline{\nu}_\mu \text{ versus } \pi^- \mu^+ \underline{\nu}_\mu$$

If nature were perfectly matter-antimatter symmetric, the rates of the above decays would be exactly equal. In fact, the rates differ by 0.3%. Our experiment eliminated the possibility that CP-violation in the two-pion decay could be some unexpected effect of the then-still-mysterious strong interactions.

Our current understanding attributes CP-violation in kaon decays not to an asymmetry in the decay process itself, but rather to an asymmetry of unknown origin in the eigenstates. Some models redefine the decay eigenstates to be:

$$|K_s\rangle = \{(1+\varepsilon)|K^0\rangle - (1-\varepsilon)|\bar{K}^0\rangle\} / \sqrt{2}$$

$$|K_L\rangle = \{(1+\varepsilon)|K^0\rangle + (1-\varepsilon)|\bar{K}^0\rangle\} / \sqrt{2}$$

$$\alpha |K_s\rangle = |K_1\rangle + \varepsilon |K_2\rangle$$

$$\alpha |K_L\rangle = |K_2\rangle + \varepsilon |K_1\rangle$$

Here  $\alpha^2 = (1+\varepsilon^*\varepsilon)$ , and  $\varepsilon$  is a complex number.

The  $|K_s\rangle$  and  $|K_L\rangle$  are “stationary” states in the sense that they decay while maintaining a constant ratio of basis state amplitudes. They are called the short-lived  $|K_s\rangle$  and long-lived  $|K_L\rangle$  neutral kaons, or more briefly, the *K-short* and the *K-long*. Adding the complex number  $\varepsilon$  means neither  $|K_s\rangle$  nor  $|K_L\rangle$  are states of definite CP, and that the  $|K_L\rangle$  has a small but non-zero amplitude to decay to two pions.

The measured value of  $\varepsilon$  is:

$$\varepsilon = (2.228 \pm 0.011) \times 10^{-3} \exp\{i (43.5 \pm 0.5)^\circ\}$$

When Madame Wu published her famous discovery of parity violation in 1956 (see ***Feynman Simplified 1D*** Chapter 49), physicists were shocked. Yet, there was some consolation in the fact that the parity violation was a 100% effect — the weak force violated parity completely. Nature was surprising, but decisive. By comparison, CP violation is even more distressing: why is nature 99.7% of what we had expected?

Fifty years later, we still have no comprehensive explanation for the origin of this tiny asymmetry.

## Self-Interaction Issues

This section addresses an advanced issue that some may consider overly technical.

Some editions of the ***Feynman Lectures*** include the following self-interaction processes in the analysis of the neutral kaon system:

$$K^0 \longrightarrow 2 \text{ pions} \longrightarrow K^0$$

$$\bar{K}^0 \longrightarrow 2 \text{ pions} \longrightarrow \bar{K}^0$$

Such reactions can be safely included here due to the feebleness of the weak force — the magnitudes of the transition amplitudes are about  $10^{-15}$ . However, self-interactions involving other forces require extreme finesse. Consider, for example, this electromagnetic self-interaction:

$$e^- \longrightarrow e^- + \text{photon} \longrightarrow e^-$$

Summing all the Feynman diagrams for all combinations of such reactions yields infinities that seemingly give the electron an infinite mass and an infinite charge, which is clearly nonsense.

This isn't a strictly quantum problem; related problems arise in classical physics as well. The potential energy of charge  $q$  in the electric field of charge  $Q$  is proportional to  $-qQ/r$ , where  $r$  is the distance between the charges. Classically, electrons are point-particles of zero size. If one tries to calculate the self-energy of one electron, the energy of its charge  $q$  in the electric field created by its charge  $q$ , the result is  $-q^2/0 = -\infty$ .

Calculations of self-interactions are explosive, in general, which is a clear indication of our incomplete understanding of these effects. We seem to have about as many problems with zero as we have with infinity.

One of Feynman's achievements that made the theory of quantum electrodynamics (QED) successful was eliminating these infinities through *renormalization*. To renormalize QED, one assumes that, in the absence of any interactions, an electron has an infinite *bare mass* and an infinite *bare charge*. These *bare* infinities are arranged to cancel the self-interaction infinities, yielding finite residuals that match the observed electron mass and charge. You might imagine what mathematicians think of that. While he didn't win the Field's Medal, Feynman did receive the Nobel Prize in Physics because his calculations match nature extremely well, which is what physicists care most about. The predictions of QED have been confirmed to as many as 12 decimal digits, and none of its predictions have ever been falsified by experiments. In 1999, Gerald 't Hooft received the Nobel Prize for refining the renormalization procedure, making it mathematical rigorous (at least in the opinion of physicists).

## Chapter 26 Review: Key Ideas

The following conservation law applies to each of the twelve types of elementary fermions:

In all processes driven by either the strong force or the electromagnetic force, the total number of fermions minus antifermions of each specific type is always conserved.

The number of bosons is not conserved, neither within one type nor summed over all types.

Only the weak force is known to change one type of elementary fermion into another, in the manner illustrated in Figure 26-1. The weak force can change some types of quarks into others, but it does not change the total number of quarks. The weak force can also change some types of leptons into others,

but it does not change the total number of leptons.

There are four kaons. The  $K^+$  and  $K^-$  are antiparticles of one another. Similarly, the  $K^0$  and  $\bar{K}^0$  (“K-zero-bar”) are different particles that are antiparticles of one another, because they contain opposite quarks. The  $K^0$  and  $\bar{K}^0$  are not two independent particles, but rather the basis states of one combined decaying two-state system. This was the first discovered example of neutral particle oscillations. To accommodate decay, the Hamiltonian of this two-state system has complex components:

$$\begin{aligned} H_{++} &= E_0 - i\beta \\ H_{+-} &= \langle K^0 | \underline{K}^0 \rangle = a + i\alpha \\ H_{-+} &= \langle \underline{K}^0 | K^0 \rangle = a + i\alpha \\ H_{--} &= E_0 - i\beta \end{aligned}$$

Absent CP-violation, the neutral kaon “stationary” states are:

$$\begin{aligned} |K_1\rangle &= (|K^0\rangle - |\underline{K}^0\rangle)/\sqrt{2} \\ |K_2\rangle &= (|K^0\rangle + |\underline{K}^0\rangle)/\sqrt{2} \end{aligned}$$

The mass difference  $m_{K_2} - m_{K_1}$  is  $3.5 \times 10^{-12}$  MeV.

This system has three time constants:

$$\begin{aligned} K_1 \text{ lifetime } t_1 &= 8.954 \times 10^{-11} \text{ sec} \\ K_2 \text{ lifetime } t_2 &= 5.116 \times 10^{-8} \text{ sec} \\ \text{Oscillation time } t_{\Delta m} &= 1.88 \times 10^{-10} \text{ sec} \end{aligned}$$

A  $K_2$  beam traversing matter *regenerates* a coherent  $K_1$  beam, because normal matter preferentially absorbs  $\underline{K}^0$  over  $K^0$ .

Because CP-violation is observed in neutral kaon decays, the decay eigenstates are the short-lived neutral kaon  $|K_S\rangle$  and the long-lived neutral kaon  $|K_L\rangle$  defined as:

$$\begin{aligned} \alpha |K_S\rangle &= |K_1\rangle + \varepsilon |K_2\rangle \\ \alpha |K_L\rangle &= |K_2\rangle + \varepsilon |K_1\rangle \end{aligned}$$

Here  $\alpha^2 = (1 + \varepsilon^* \varepsilon)$ , and  $\varepsilon$  is a complex number measured to be:  
 $\varepsilon = (2.228 \pm 0.011) \times 10^{-3} \exp\{i (43.5 \pm 0.5)^\circ\}$

Parity violation in weak interactions is a 100% effect. By comparison, CP violation is only a 0.3% effect. Fifty years later, we still have no comprehensive explanation for the origin of this tiny asymmetry.

# Chapter 27

## Angular Momentum

---

In Chapter 23, we discovered that the conservation of angular momentum is linked in quantum mechanics to the symmetry of natural laws under rotation. Recall how quantum state  $|\psi\rangle$  is changed by a rotation of the coordinate system by angle  $\theta$  about the z-axis:

$$R_z(\theta) |\psi\rangle = \exp\{iJ_z\theta/\hbar\} |\psi\rangle$$

Here  $J_z$  is the z-component of angular momentum of state  $|\psi\rangle$ . If natural laws are symmetric under rotation about the z-axis,  $R_z(\theta)$  commutes with any Hamiltonian  $H$ , and the quantity  $J_z$  is conserved.

In Chapter 25, we examined how angular momentum conservation determines the angular distribution of protons coming from decaying  $\Lambda^0$ s.

In this chapter, we will explore the role of angular momentum conservation in other atomic systems.

### Radiation by Excited Atoms

We know that electrons in atoms are confined to specific orbits characterized by three quantum numbers:  $n$ ,  $j$ , and  $m$ . The principal quantum number  $n$  determines the orbits' average inverse radius  $\langle 1/r \rangle$ . Quantum number  $j$  specifies the orbit's total angular momentum, and  $m$  specifies the component of angular momentum along any selected axis (we usually select the z-axis).

The second law of thermodynamics favors processes that convert potential energy into kinetic energy. Because kinetic energy can be distributed in more ways than can potential energy, the same amount of energy has more entropy if it's kinetic rather than potential. Recall from ***Feynman Simplified 1B*** Chapter 23 that a system's entropy is proportional to the logarithm of the number of microscopically distinct, but macroscopically indistinguishable, internal arrangements of that system.

No form of energy has more entropy than radiation. This is why an electron in a higher energy atomic orbit, an *excited state*, is driven to emit radiation and drop to a vacancy in a lower energy orbit.

Let's examine how an emitted photon is affected by an electron's initial angular momentum.

We start with the simplest case. Consider an atom in an excited state with angular momentum parallel to the z-axis that emits a photon in the +z-direction. Two scenarios are shown in Figure 27-1: an initial excited state  $|A\rangle$  with  $j=1$  and  $m=+1$ ; and an initial excited state  $|C\rangle$  with  $j=1$  and  $m=-1$ .

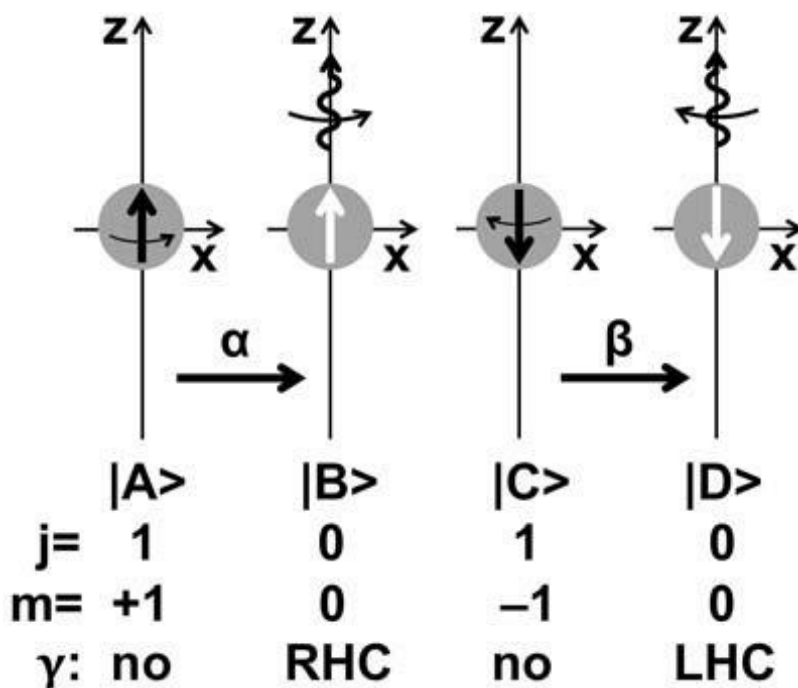


Figure 27-1  
States  $|A\rangle$  &  $|C\rangle$  Emitting Photons

In Figure 27-1, state  $|A\rangle$  transitions to state  $|B\rangle$  with amplitude  $\alpha$ , and state  $|C\rangle$  transitions to state  $|D\rangle$  with amplitude  $\beta$ . Each gray disk represents an atom. The bold arrows in  $|A\rangle$  and  $|C\rangle$  represent the atom's initial angular momentum, while the white arrows in  $|B\rangle$  and  $|D\rangle$  are reminders of what that angular momentum was originally. The wavy lines represent emitted photons.

To conserve angular momentum along the z-axis, the photon in state  $|B\rangle$  must have spin +1, and the photon in state  $|D\rangle$  must have spin -1. As we discovered in Chapter 23, this means the photon in  $|B\rangle$  has RHC (right hand circular) polarization, and the photon in  $|D\rangle$  has LHC (left hand circular) polarization.

In V3p18-2, Feynman notes that to be more precise, we should say  $\alpha$  and  $\beta$  are the amplitudes for photons to be emitted within an infinitesimal solid angle  $d\Omega$  surrounding the +z-axis during an infinitesimal time interval  $dt$ . We write this as:

$$\langle B|A\rangle d\Omega dt = \alpha d\Omega dt$$

$$\langle D|C\rangle d\Omega dt = \beta d\Omega dt$$

Since a photon cannot have spin 0 along any axis, a photon cannot be emitted parallel to the z-axis by an atom with  $m=0$  along that axis.

We will now proceed to relate amplitudes  $\alpha$  and  $\beta$  by employing inversion and reflection symmetries. Feynman says rotations, translations, and inversions should generally be performed by changing the coordinate axes rather than by changing the objects being considered. However in this case, it will be much easier to see what is happening if we change the objects rather than the coordinate axes.

Let's first recall how axial and polar vectors respond to inversion (see *Feynman Simplified 1D* Chapter 49). When an object is inverted, we reverse the polarity of the x, y, and z coordinates of each piece of the object. In this process, each component of every polar vector reverses polarity, but no component of any axial vector changes. Polar vectors include position, velocity and momentum. Axial

vectors include angular momentum and torque. Recall that angular momentum  $\mathbf{L}$  equals  $\mathbf{r} \times \mathbf{p}$ ; under inversion, both  $\mathbf{r}$  and  $\mathbf{p}$  reverse polarity and so  $\mathbf{L}$  does not.

We will now invert and then rotate state  $|B\rangle$ , and find that the result is state  $|D\rangle$ .

Consider Figure 27-2, in which states  $|A\rangle$  and  $|B\rangle$  are exactly as above.

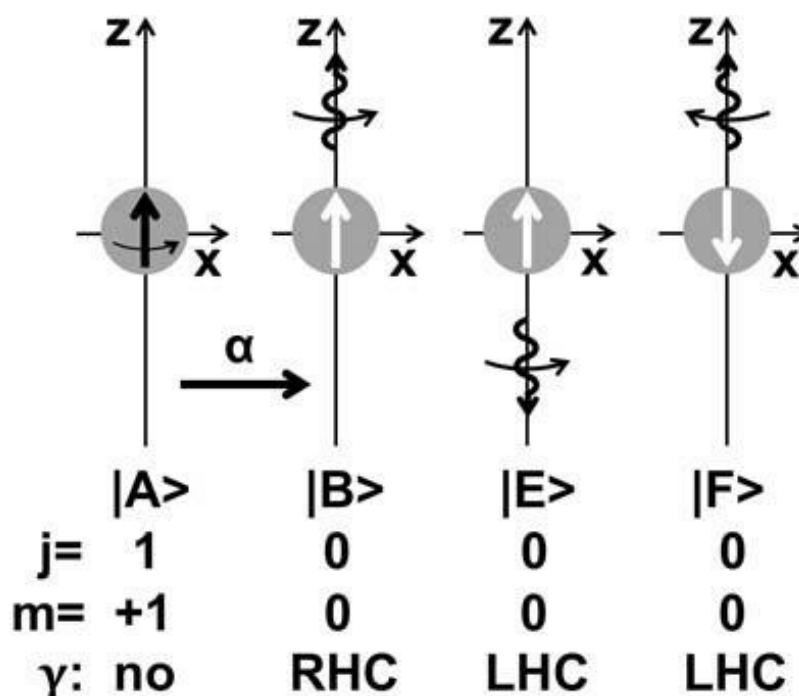


Figure 27-2  
Invert & Rotate  $|B\rangle$  To Get  $|F\rangle$

State  $|E\rangle$  is state  $|B\rangle$  inverted through the atom's center. In  $|E\rangle$ , the photon is moving toward  $-z$ . The atom's original angular momentum (white arrow) is an axial vector that does not change under inversion. Similarly, the photon's angular momentum does not change direction; it is still pointing upward, but referenced to its motion toward  $-z$  that now has LHC polarization.

State  $|F\rangle$  is state  $|E\rangle$  rotated by  $180^\circ$  about the  $y$ -axis that points into the screen. This rotation turns the atom's original angular momentum (white arrow) upside-down. The photon retains LHC polarization because rotation does not change its spin-velocity relationship.

Comparing Figures 27-1 and 27-2, we see that state  $|D\rangle$  looks identical to state  $|F\rangle$ . In quantum mechanics, this means the two states can differ by at most a phase angle. We can further constrain this phase angle by performing a second inversion and rotation. Two inversions are equivalent to a null operation. As we discovered studying the rotation of bosons and fermions, two  $180^\circ$  rotations can change an amplitude only by a factor of  $\pm 1$ . All this means:

$$\alpha = \pm \beta$$

## Parity Considerations

We know that parity is conserved in electromagnetic interactions. Let's see how parity conservation affects the emission process we are discussing.

The states examined above are not states of definite parity, but we can construct combinations of those photon states that do have definite parity. Recall that we obtained state  $|E\rangle$  by inverting state  $|B\rangle$ . Both states involve an atom with  $m=+1$  along the z-axis. In  $|B\rangle$ , the photon moves toward  $+z$  with RHC polarization; relabel this state  $|RHC,+z\rangle$ . In  $|E\rangle$ , the photon moves toward  $-z$  with LHC polarization; relabel this state  $|LHC,-z\rangle$ . With  $P$  being the parity operator for inverting objects (or coordinate systems), this means:

$$\begin{aligned} |LHC,-z\rangle &= P |RHC,+z\rangle \\ |RHC,+z\rangle &= P |LHC,-z\rangle \end{aligned}$$

Now define  $|\psi+\rangle$  as the state of any photon emitted parallel to the z-axis by an atom with  $m=+1$ . Express  $|\psi+\rangle$  as a linear combination of  $|RHC,+z\rangle$  and  $|LHC,-z\rangle$  with unknown coefficients  $\lambda$  and  $\mu$ .

$$\begin{aligned} |\psi+\rangle &= \lambda |RHC,+z\rangle + \mu |LHC,-z\rangle \\ P|\psi+\rangle &= \lambda |LHC,-z\rangle + \mu |RHC,+z\rangle \end{aligned}$$

For  $|\psi+\rangle$  to be a photon state of definite parity, we must have:

$$P |\psi+\rangle = \pm |\psi+\rangle$$

This requires:

$$\begin{aligned} \mu &= \pm \lambda = 1/\sqrt{2} \\ |\psi+\rangle &= ( |RHC,+z\rangle \pm |LHC,-z\rangle )/\sqrt{2} \end{aligned}$$

Here, the plus sign corresponds to even parity and the minus sign to odd parity.

The parity of the initial and final atomic states could be the same or could be the opposite. That results in two distinct situations with two distinct photon angular distributions. Let's examine both possibilities.

The first possibility is that the initial and final atomic states have opposite parity. Here, the photon state must have odd parity for the entire atom-photon system to conserve parity. This means:

$$|\psi+\rangle = ( |RHC,+z\rangle - |LHC,-z\rangle )/\sqrt{2}$$

Let's rotate this state by  $180^\circ$  about the y-axis. As we discovered above, this reverses the polarity of  $m$ , and the photon direction but not its polarization. Changing  $m$  from  $+1$  to  $-1$  changes  $|\psi+\rangle$  into  $|\psi-\rangle$ .

According to the rotation tables:

$$\begin{aligned} \langle m=-1 | R_y(180^\circ) | m=+1 \rangle &= 1 \\ |\psi-\rangle &= R_y(180^\circ) |\psi+\rangle \\ |\psi-\rangle &= ( |RHC,-z\rangle - |LHC,+z\rangle )/\sqrt{2} \end{aligned}$$

We highlight two amplitudes from the prior four equations:

$$\langle \psi+ | \text{RHC}, +z \rangle = +1/\sqrt{2}$$

$$\langle \psi- | \text{LHC}, +z \rangle = -1/\sqrt{2}$$

The upper line corresponds to amplitude  $\alpha$  from the prior section: the amplitude for an  $m=+1$  state to emit a photon moving toward  $+z$  with RHC polarization. The lower line corresponds to amplitude  $\beta$  from the prior section: the amplitude for an  $m=-1$  state to emit a photon moving toward  $+z$  with LHC polarization. This proves that:

$$\alpha = -\beta, \text{ for a photon state of odd parity}$$

If the initial and final atomic states have opposite parity, which is very common, the transition is said to produce *electric dipole radiation*.

The second possibility is that the initial and final atomic states have the same parity. Here, the photon state must have even parity for the entire atom-photon system to conserve parity. This transition is said to produce *magnetic dipole radiation*. The photon state is then:

$$|\psi+\rangle = ( |\text{RHC}+z\rangle + |\text{LHC}-z\rangle )/\sqrt{2}$$

Rotating by  $180^\circ$  about the y-axis yields:

$$|\psi-\rangle = R_y(180^\circ) |\psi+\rangle$$

$$|\psi-\rangle = ( |\text{RHC}-z\rangle + |\text{LHC}+z\rangle )/\sqrt{2}$$

$$\langle \psi+ | \text{RHC}+z \rangle = +1/\sqrt{2}$$

$$\langle \psi- | \text{LHC}+z \rangle = +1/\sqrt{2}$$

$$\alpha = +\beta, \text{ for a photon state of even parity}$$

We can now calculate the amplitude for a photon to be emitted in any direction. Define the photon direction to be the  $z^*$ -axis. Define the  $z$ -axis to be the direction of the atom's angular momentum, so that  $m=+1$  in the  $z$ -direction. Let  $\theta$  be the angle between the  $z$ - and  $z^*$ -axes, and define the  $y$ -axis such that  $z^*$  is obtained by rotating the  $z$ -axis by angle  $\theta$  about the  $y$ -axis.

We can use the rotation tables for  $R_y(\theta)$ , listed at the end of Chapter 17, to project the state  $m=+1$  along the  $z$ -axis onto the states  $m=+1$ ,  $0$ , and  $-1$  along the  $z^*$ -axis. For example:

$$\langle +1, z^* | R_y(\theta) | +1, z \rangle = (1+\cos\theta)/2$$

The amplitude to emit a RHC photon along the  $z^*$ -axis equals the product of (amplitude of RHC parallel to  $m$ )  $\times$  (amplitude of  $m=+1$  along  $z^*$ -axis), which is:

$$\text{For RHC: } (\alpha) (\langle +1 | R_y(\theta) | +1 \rangle) = \alpha (1+\cos\theta)/2$$

The amplitude to emit a LHC photon along the  $z^*$ -axis equals the product of (amplitude of LHC antiparallel to  $m$ )  $\times$  (amplitude of  $m=-1$  along  $z^*$ -axis), which is:

For LHC:  $(\beta) \langle -1 | R_y(\theta) | +1 \rangle = \beta (1 - \cos\theta)/2$

Recall that for an odd parity photon state  $\beta = -\alpha$ , and for an even parity state  $\beta = +\alpha$ .

Appropriate combinations of the above provide the amplitudes for x- and y-polarizations.

## Light Scattering

In V3p18-3, Feynman applies the results of the prior section to a “somewhat more complicated problem—but also one that is somewhat more real”: the scattering of light by atoms.

We assume an atom is in its ground state with  $j=0$  and  $m=0$ . Assume also that a RHC polarized photon approaching from the  $-z$ -direction is absorbed by the atom, and then subsequently re-emitted with RHC polarization at angle  $\theta$  relative to the  $z$ -axis. This two-step process is illustrated in Figure 27-3.

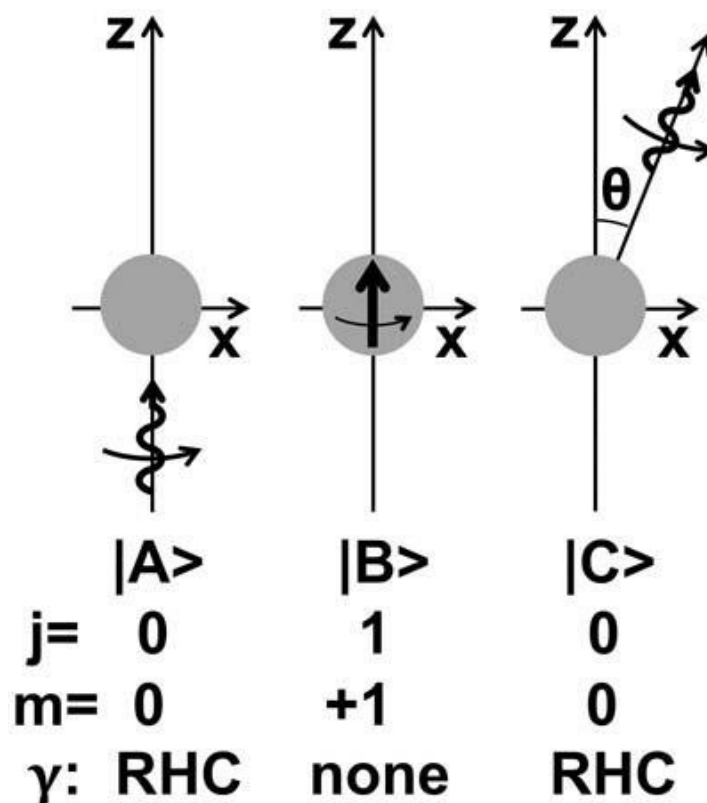


Figure 27-3  
Light Scattering

In the first step, going from state  $|A\rangle$  to state  $|B\rangle$ , the conservation of angular momentum requires the atom to enter its  $j=1$ ,  $m=+1$  state after absorbing the photon. Call the amplitude for this process  $\mu$ .

In the second step, going from  $|B\rangle$  to  $|C\rangle$ , the photon is emitted with RHC polarization with the amplitude we calculated above:

$$\alpha (1 + \cos\theta)/2$$

The amplitude for the total process is the product of the amplitudes for each step, which we label as follows:

$$\langle R|\hat{S}|R\rangle = \alpha \mu (1+\cos\theta)/2$$

Here  $\hat{S}$  is the scattering operator. The left side of the equation reads: a RHC polarized photon is scattered, resulting in a RHC photon emitted at angle  $\theta$ .

The atom could also emit a LHC polarized photon at angle  $\theta$ . If we assume the photon state has odd parity ( $\beta=-\alpha$ ), the amplitude for this total process is:

$$\langle L|\hat{S}|R\rangle = -\alpha \mu (1-\cos\theta)/2$$

We could repeat this analysis for an incident LHC polarized photon, which drives the absorbing atom into the  $m=-1$  state. By the same logic employed in the prior section, and again assuming odd parity, the amplitude for the atom to absorb a LHC photon and enter the  $m=-1$  state equals  $-\mu$ .

The amplitudes for the total processes for a LHC incident photon are:

$$\langle R|\hat{S}|L\rangle = -\alpha \mu (1-\cos\theta)/2$$

$$\langle L|\hat{S}|L\rangle = (-\alpha)(-\mu)(1+\cos\theta)/2$$

Now, examine what happens if the incident photon has +x-polarization.

The principle here is simple enough: x-polarization is a superposition of RHC and LHC. We know the amplitudes for RHC and LHC, hence we combine these amplitudes using the superposition coefficients.

From Chapter 15, the superposition coefficients for x- and y-polarizations are:

$$|x\rangle = (|RHC\rangle + |LHC\rangle)/\sqrt{2}$$

$$|y\rangle = (|RHC\rangle - |LHC\rangle)/(i\sqrt{2})$$

Let's take one example: an incident x-polarized photon, with an emitted RHC photon at angle  $\theta$ .

$$\langle R|\hat{S}|x\rangle = \{\langle R|\hat{S}|R\rangle + \langle R|\hat{S}|L\rangle\}/\sqrt{2}$$

$$\langle R|\hat{S}|x\rangle = \alpha \mu \{(1+\cos\theta) - (1-\cos\theta)\}/(2\sqrt{2})$$

$$\langle R|\hat{S}|x\rangle = \alpha \mu \cos\theta/\sqrt{2}$$

For an emitted LHC photon at angle  $\theta$  we have:

$$\langle L|\hat{S}|x\rangle = \{\langle L|\hat{S}|R\rangle + \langle L|\hat{S}|L\rangle\}/\sqrt{2}$$

$$\langle L|\hat{S}|x\rangle = \alpha \mu \{-(1-\cos\theta) + (1+\cos\theta)\}/(2\sqrt{2})$$

$$\langle L|\hat{S}|x\rangle = \alpha \mu \cos\theta/\sqrt{2}$$

Reversing the polarization relationships yields:

$$|RHC\rangle = (|x\rangle + i|y\rangle)/\sqrt{2}$$

$$|LHC\rangle = (|x\rangle - i|y\rangle)/\sqrt{2}$$

$$\langle x|RHC\rangle = 1/\sqrt{2}$$

$$\langle x|LHC\rangle = 1/\sqrt{2}$$

$$\langle y|RHC\rangle = i/\sqrt{2}$$

$$\langle y|LHC\rangle = -i/\sqrt{2}$$

From these amplitudes we calculate the amplitude that the emitted photon has x-polarization.

$$\begin{aligned}\langle x@ \theta |S|x\rangle \\ = \{ \langle x|R\rangle \langle R@ \theta |S|x\rangle + \langle x|L\rangle \langle L@ \theta |S|x\rangle \} / \sqrt{2}\end{aligned}$$

$$\langle x@ \theta |S|x\rangle = (\alpha \mu \cos \theta) / 2 + (\alpha \mu \cos \theta) / 2$$

$$\langle x@ \theta |S|x\rangle = \alpha \mu \cos \theta$$

The amplitude that the emitted photon has y-polarization is:

$$\begin{aligned}\langle y@ \theta |S|x\rangle \\ = \{ \langle y|R\rangle \langle R@ \theta |S|x\rangle + \langle y|L\rangle \langle L@ \theta |S|x\rangle \} / \sqrt{2}\end{aligned}$$

$$\langle y@ \theta |S|x\rangle = (i \alpha \mu \cos \theta) / 2 + (-i \alpha \mu \cos \theta) / 2$$

$$\langle y@ \theta |S|x\rangle = 0$$

The last equation means that, for plane polarization, the scattered light is polarized in the same plane as the incident light.

In V3p18-4, Feynman remarks that the results just obtained were also obtained in our classical analysis of light scattering (*Feynman Simplified 1C* Chapter 35). There we assumed atoms acted like harmonic oscillators.

Feynman says:

“Perhaps you are thinking: ‘It’s so much easier in the classical theory; if it gives the right answer why bother with the quantum theory?’ For one thing, we have considered so far only the special—though common—case of an atom with a  $j=1$  excited state and a  $j=0$  ground state. If the excited state had spin two, you would get a different result. Also, there is no reason why the model of an electron attached to a spring and driven by an oscillating electric field should work for a single photon. But we have found that it does in fact work, and that the polarization and intensities come out right. So in a certain sense we are bringing the whole course around to the real truth. Whereas we have, in Vol. I, done the theory of the index of refraction, and of light scattering, by the classical theory, we have now shown that the quantum theory gives the same result for the most common case. In effect we have now done the polarization of sky light, for instance, by quantum mechanical arguments, which is the only truly legitimate way.

“It should be, of course, that all the classical theories which work are supported ultimately by legitimate quantum arguments. Naturally, those things which we have spent a great deal of time in explaining to you were selected from just those parts of classical physics which still maintain

validity in quantum mechanics.”

## Positronium Annihilation

Positronium is a bound state of an electron and an antielectron — a positron. Being half matter and half antimatter, it is not surprising that positronium is relatively short-lived. It could be described as an “atom” with a short half-life.

Feynman says positronium is “very pretty” and “quite interesting and, although somewhat complicated, we hope not too much so.”

Like hydrogen, positronium has excited states, but we will examine only its ground state, which has orbital angular momentum  $j=0$ . Also, like hydrogen, positronium contains two spin  $1/2$  fermions, whose spin-spin interactions cause *hyperfine* splitting of the ground state. As we discovered in Chapter 16, this results in a spin-zero singlet state and a spin-one triplet state. In the absence of external magnetic fields, the three triplet substates all have the same energy, which is higher than the energy of the singlet state.

Because the positron’s magnetic dipole moment is 1000 times larger than the proton’s, the hyperfine splitting in positronium is about 1000 times larger than in hydrogen.

When the positron and the electron annihilate, their mass energy is converted into the kinetic energy of two or more photons. Subatomic annihilation has no macroscopic analog — it is absolute destruction that leaves no trace whatsoever of the original particles.

Let’s first consider the singlet state. As in hydrogen, the singlet state  $|j,m\rangle = |0,0\rangle$  is an antisymmetric superposition of antiparallel spins, represented by:

$$|0,0\rangle = \{|e^+,p^-\rangle - |e^-,p^+\rangle\}/\sqrt{2}$$

Here  $e^+$  denotes an electron with spin  $+1/2$ ,  $p^-$  denotes a positron with spin  $-1/2$ , etc. In its rest frame, the singlet state annihilates with a mean lifetime of  $1.25 \times 10^{-10}$  seconds into two 511 keV photons. These photons, in the gamma ray energy range, must have exactly the same energy and exactly opposite momenta in the CM frame. Since the singlet state has spin zero, it has no spin axis or other preferred direction. This means all photon directions are equally probable, their angular distribution is uniform in the CM frame.

Now let’s consider the polarization of the photons emitted in the disintegration of singlet state positronium, as illustrated in Figure 27-4.

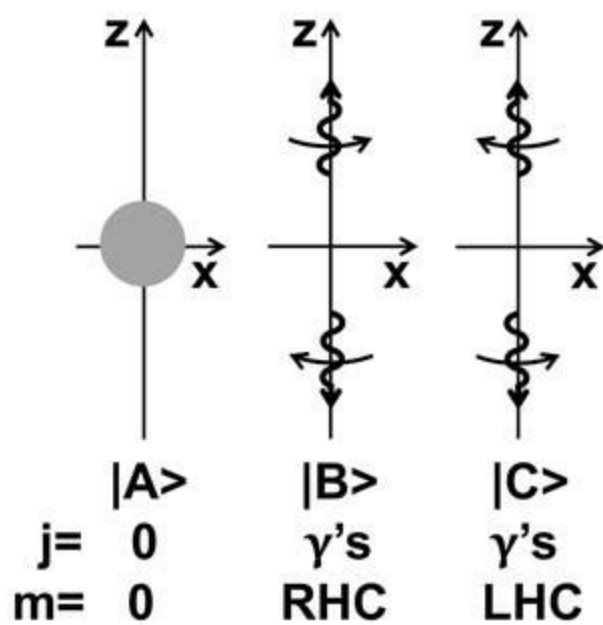


Figure 27-4  
Decay of  $e^+e^-$  to 2 Photons

We may describe photon polarization in various ways, but RHC and LHC polarizations are the easiest to relate to angular momentum. Since the initial state  $|A\rangle$  has zero angular momentum, angular momentum conservation requires that the two photons in the final state have equal but opposite angular momenta. This means they must both have the same circular polarization: either RHC, as in  $|B\rangle$ ; or LHC, as in  $|C\rangle$ . With both photons RHC polarized, the upper photon in  $|B\rangle$  has spin  $+1$  along its direction of motion  $+z$ , while the lower photon has spin  $+1$  along its direction of motion  $-z$ , which is the same as spin  $-1$  in the  $+z$ -direction.

We will get back to the singlet state shortly, but let's first consider the triplet state. It has angular momentum  $j=1$ , and three substates with  $m=+1$ ,  $0$ , and  $-1$ . The  $m=\pm 1$  substates cannot decay into two photons, because two spin 1 photons can only have total spin  $+2$ ,  $0$ , or  $-2$  in any direction. (If that isn't clear, compute the possible sums of  $\pm 1$  and  $\pm 1$ . Photons can never have spin 0 along any axis.)

What about state  $|10\rangle$ , the state with  $j=1$  and  $m=0$ ? As it turns out, a two-photon decay is also impossible for  $|10\rangle$ . Let's see why.

Consider a two-photon final state, such as  $|B\rangle$  in figure 27-4. Rotating  $|B\rangle$  by  $180^\circ$  about the  $y$ -axis results in the same state that we started with; it simply exchanges the two photons. Being bosons, we know that the transition amplitudes to the original  $|B\rangle$  and to the photon-exchanged  $|B\rangle$  must be equal:

$$R_y(180^\circ) \langle B|10\rangle = + \langle B|10\rangle$$

However, rotating  $|10\rangle$  by  $180^\circ$  about the  $y$ -axis multiplies that state by  $\cos(180^\circ)$ , according to the Chapter 17 rotation tables for  $j=1$ ,  $m=0$ . This means the amplitudes to transition from the original  $|10\rangle$  and the rotated  $|10\rangle$  must have opposite signs:

$$R_y(180^\circ) \langle B|10\rangle = - \langle B|10\rangle$$

The only solution to both requirements is  $\langle B|10\rangle=0$ . This means the spin-one triplet state must decay to three photons, rather than two. The three-photon decay rate is less than the two-photon rate, which

extends the triplet state lifetime to  $1.42 \times 10^{-7}$  seconds, over 1000 times longer than the singlet state lifetime.

Angular momentum and parity can also be conserved by: the triplet state decaying to any odd number of photons greater than three; or by the singlet state decaying to any even number of photons greater than two. But, the decay rates drop by a factor of one million for each additional pairs of photons.

When positronium forms, the two spin orientations are typically random, resulting in the longer-lived triplet state 75% of the time and the shorter-lived singlet state 25% of the time.

Now, we return to the singlet state two-photon decay shown in Figure 27-4. We learned above that, due to angular momentum conservation, the decay must result in the following two states (or a linear combination thereof):

$|B\rangle$ : two RHC polarized photons

$|C\rangle$ : two LHC polarized photons

Because parity is conserved in electromagnetic processes, the actual state must have the same parity as positronium.

Let's consider how parity acts on these states. In the analysis of Figure 27-2, we found that  $P$ , the parity/spatial-inversion operator, acting on a RHC polarized photon reverses its direction of motion and also changes its polarization to LHC. Similarly  $P$  acting on a LHC photon reverses its direction of motion and changes its polarization to RHC. Because, in the positronium rest frame, both  $|B\rangle$  and  $|C\rangle$  have two photons moving in opposite directions, reversing their directions has no effect. The only effect of  $P$  operating on  $|B\rangle$  and  $|C\rangle$  is to invert their polarizations, which means exchanging  $|B\rangle$  and  $|C\rangle$ .

$$P |B\rangle = |C\rangle$$

$$P |C\rangle = |B\rangle$$

This means neither state has definite parity. The actual final state of positronium decay must be a linear superposition of  $|B\rangle$  and  $|C\rangle$ .

Now what is the parity of positronium? Advanced quantum theory, confirmed by experiment, shows that any elementary fermion and its antiparticle have opposite parity. (Elementary bosons are their own antiparticles, so this logic does not apply to bosons.) Theory and experiment also show that the parity of a composite system, such as positronium, equals the product of the parities of all its constituents multiplied by  $(-1)^j$ , where  $j$  is the system's total angular momentum.

Hence, the parity of the singlet ( $j=0$ ) state of positronium is:

$$(\text{parity of } e^-) \times (\text{parity of } e^+) \times (-1)^0 = -1$$

A linear superposition of  $|B\rangle$  and  $|C\rangle$  that has parity  $-1$  (odd parity) is:

$$|F\rangle = (|B\rangle - |C\rangle)/\sqrt{2}$$

Let's confirm the parity of  $|F\rangle$ .

$$P|F\rangle = (P|B\rangle - P|C\rangle)/\sqrt{2}$$

$$P|F\rangle = (|C\rangle - |B\rangle)/\sqrt{2} = -|F\rangle$$

For the decay of the singlet state of positronium, the only final state that conserves both angular momentum and parity is  $|F\rangle$ , a spin zero, antisymmetric superposition of a RHC-pair and a LHC-pair of photons moving in opposite directions in the positronium rest frame.

## Entanglement

The photons from positronium decay are a prime example of *entanglement*, a uniquely quantum phenomenon. The two photons are not independent particles, each with its own wave function. Instead, the two photons are entangled, sharing one common wave function.

In V3p18-7, Feynman highlights a stunning consequence of entanglement. Imagine having a source at  $x=y=z=0$  that produces one singlet-state positronium “atom” each second, with the atom at rest in the lab frame. Now imagine scientist Alice monitoring a photon detector at  $x=y=0$ ,  $z=+D$ , and scientist Bob monitoring a photon detector at  $x=y=0$ ,  $z=-D$ . Whenever a photon from positronium annihilation hits Alice's detector, its partner must hit Bob's detector, because the annihilation photons must move in exactly opposite directions in the positronium rest frame. Now imagine that each detector flashes a red light when a RHC photon arrives and a lavender light when a LHC photon arrives.

Alice and Bob both see red and lavender with equal probability, because state  $|F\rangle$  is an equal superposition of RHC and LHC. Neither can predict what color the next light will be, but the odds are 50-50, just as though each scientist was flipping a perfect coin. The interesting thing is: they are flipping the same coin. Whenever Alice sees red so must Bob; when she sees lavender so does he. It **never** happens that one sees red when the other sees lavender, even if they are light-years apart.

This begs the question: If Alice's photon has an equal chance of being red or lavender, how does Bob's photon “know” what Alice's photon is going to be? This question launched a vigorous debate that continued for many decades. We will come back to this “Einstein-Podolsky-Rosen (EPR) paradox” in Chapters 34 and 35.

There is one last twist to the positronium story.

Imagine now that Alice and Bob change their detectors to trigger on x- and y-polarizations instead of RHC and LHC polarizations. Recall the quantum mechanical relationships between these states.

$$|RHC\rangle = (|x\rangle + i|y\rangle)/\sqrt{2}$$

$$|LHC\rangle = (|x\rangle - i|y\rangle)/\sqrt{2}$$

$$|x\rangle = (|RHC\rangle + |LHC\rangle)/\sqrt{2}$$

$$|y\rangle = (|RHC\rangle - |LHC\rangle)/(i\sqrt{2})$$

The photons have amplitude  $1/\sqrt{2}$  to be RHC, in which case the amplitude to have x-polarization is  $1/\sqrt{2}$  and the amplitude to have y-polarization is  $i/\sqrt{2}$ . The photons also have amplitude  $-1/\sqrt{2}$  to be LHC, in which case the amplitude to have x-polarization is  $1/\sqrt{2}$  and the amplitude to have y-polarization is  $-i/\sqrt{2}$ .

Therefore, the amplitude for Alice to detect x-polarization when Bob detects y-polarization is:

$$\begin{aligned} \langle A_x B_y | F \rangle &= \langle A_x | \text{RHC} \rangle \langle B_y | \text{RHC} \rangle \langle \text{RHC} | F \rangle \\ &+ \langle A_x | \text{LHC} \rangle \langle B_y | \text{LHC} \rangle \langle \text{LHC} | F \rangle \end{aligned}$$

$$\begin{aligned} \langle A_x B_y | F \rangle &= (1/\sqrt{2}) (i/\sqrt{2}) (+1/\sqrt{2}) \\ &+ (1/\sqrt{2}) (-i/\sqrt{2}) (-1/\sqrt{2}) \end{aligned}$$

$$\langle A_x B_y | F \rangle = i/\sqrt{8} + i/\sqrt{8} = i/\sqrt{2}$$

And the amplitude for Alice to detect x-polarization when Bob detects x-polarization is:

$$\begin{aligned} \langle A_x B_x | F \rangle &= \langle A_x | \text{RHC} \rangle \langle B_x | \text{RHC} \rangle \langle \text{RHC} | F \rangle \\ &+ \langle A_x | \text{LHC} \rangle \langle B_x | \text{LHC} \rangle \langle \text{LHC} | F \rangle \end{aligned}$$

$$\begin{aligned} \langle A_x B_x | F \rangle &= (1/\sqrt{2}) (1/\sqrt{2}) (+1/\sqrt{2}) \\ &+ (1/\sqrt{2}) (1/\sqrt{2}) (-1/\sqrt{2}) \end{aligned}$$

$$\langle A_x B_x | F \rangle = 1/\sqrt{8} - 1/\sqrt{8} = 0$$

By symmetry,

$$\begin{aligned} \langle A_y B_x | F \rangle &= \langle A_x B_y | F \rangle = i/\sqrt{2} \\ \langle A_y B_y | F \rangle &= \langle A_x B_x | F \rangle = 0 \end{aligned}$$

This means the probability of the two polarizations being equal (xx or yy) is zero, and the probability of the polarizations being orthogonal (xy or yx) is 100% (50% for each of xy and yx). Note that in V3p18-8, Feynman calculates  $\langle A_x B_y | F \rangle = i$ , which would incorrectly make the total probability of perpendicular polarizations (xy and yx) equal to 200%. This error persists in even recent editions.

## Chapter 27 Review: Key Ideas

- **Parity operator P** acting on circularly polarized photons reverses the direction of motion and exchanges RHC and LHC polarizations. Since all the processes discussed in this chapter are electromagnetic, they all conserve parity.

Elementary fermions and their antiparticles have opposite parity. Bosons are their own antiparticles. The parity of a composite system equals the product of the parities of all its constituents multiplied by  $(-1)^j$ , where  $j$  is the system's total angular momentum.

- For an **atom in an excited state**, with angular momentum  $j=1$  and component of angular momentum  $m=+1$  along the  $z$ -axis, let  $\alpha$  be the amplitude to emit a RHC photon in the  $+z$ -direction, and let  $\beta$  be the amplitude to emit a LHC photon in the  $-z$ -direction. The amplitude to emit a circularly polarized photon at angle  $\theta$  relative to the  $z$ -axis is:

For RHC:  $\alpha (1+\cos\theta)/2$

For LHC:  $\beta (1-\cos\theta)/2$

Neither RHC nor LHC have definite parity. The actual photon state is a linear combination of RHC and LHC that does have definite parity. In *electric dipole radiation*, the atom's parity reverses, requiring the photon state to have odd parity and making  $\beta=-\alpha$ . The photon state is:

$$|\psi\rangle = \{(1+\cos\theta)|R, +z\rangle - (1-\cos\theta)|L, -z\rangle\}/\sqrt{8}$$

Here,  $|R, +z\rangle$  is the state of a RHC photon moving toward  $+z$ , and  $|L, -z\rangle$  is the state of a LHC photon moving toward  $-z$ .

In *magnetic dipole radiation*, the atom's parity is unchanged, requiring the photon state to have even parity and making  $\beta=+\alpha$ . The photon state is:

$$|\psi\rangle = \{(1+\cos\theta)|R, +z\rangle + (1-\cos\theta)|L, -z\rangle\}/\sqrt{8}$$

- **Light scattering** by an atom can be considered a two-step process: absorption followed by emission. We will assume both steps change the atom's parity. Let  $\mu$  be the amplitude for an atom to absorb a RHC photon and enter a  $j=1, m=+1$  state. It follows that  $-\mu$  is the amplitude to absorb a LHC photon and enter a  $j=1, m=-1$  state. With scattering operator  $S$ , the amplitudes for incident RHC or LHC photons to scatter at angle  $\theta$  with RHC or LHC polarizations are:

$$\langle R @ \theta | S | R \rangle = + \alpha \mu (1+\cos\theta)/2$$

$$\langle L @ \theta | S | R \rangle = - \alpha \mu (1-\cos\theta)/2$$

$$\langle R @ \theta | S | L \rangle = - \alpha \mu (1-\cos\theta)/2$$

$$\langle L @ \theta | S | L \rangle = + \alpha \mu (1+\cos\theta)/2$$

If the incident photon has  $+x$ -polarization, the amplitudes for the polarization states of the emitted photon are:

$$\langle R @ \theta | S | x \rangle = \alpha \mu \cos\theta/\sqrt{2}$$

$$\langle L @ \theta | S | x \rangle = \alpha \mu \cos\theta/\sqrt{2}$$

$$\langle x @ \theta | S | x \rangle = \alpha \mu \cos\theta$$

$$\langle y @ \theta | S | x \rangle = 0$$

The last equation means that, for plane polarization, the scattered light is polarized in the same plane

as the incident light.

- **Positronium** is a bound system of an electron and an antielectron (a positron). Like hydrogen, the orbital ground state has four spin states: a spin-one triplet state, with  $m=+1, 0$ , and  $-1$ ; and a spin-zero singlet state. The orbital ground state has parity  $-1$  (odd). Due to the conservation of both angular momentum and parity, only the singlet state can decay to two photons; the triplet state must decay to three photons, and therefore its lifetime is 1000 times longer ( $1.42 \times 10^{-7}$  vs.  $1.25 \times 10^{-10}$  seconds).

For the decay of the singlet state of positronium, the only final state that conserves both angular momentum and parity is a spin zero, antisymmetric superposition of a RHC-pair and a LHC-pair of photons moving in opposite directions in the positronium rest frame.

- **Entanglement** is the uniquely quantum phenomenon of two or more particles sharing one common wave function. The two photons from positronium decay are a prime example.

Scientists Alice and Bob, monitoring photon detectors on opposite sides of a singlet-state positronium annihilation source, will always see both photons with the same circular polarization, even though each has a 50-50 chance of being either RHC or LHC. It never happens that one sees RHC when the other sees LHC, even if they are light-years apart.

If Alice and Bob switch their detectors to x- and y-polarizations, the amplitudes for Alice and Bob to detect x-polarization and y-polarization are:

$$\langle A_x B_y | e^+ e^- \rangle = i/\sqrt{2}$$

$$\langle A_x B_x | e^+ e^- \rangle = 0$$

$$\langle A_y B_x | e^+ e^- \rangle = i/\sqrt{2}$$

$$\langle A_y B_y | e^+ e^- \rangle = 0$$

This means the probability of the two polarizations being equal (xx or yy) is zero, and the probability of the polarizations being orthogonal (xy or yx) is 100%.

# Chapter 28

## Rotations Of Any Spin

---

For your convenience, all rotation matrices derived here are presented at the end of this chapter.

There are three key principles when summing angular momenta from several sources.

Firstly,  $M = \sum_n m_n$ .  $M$ , the z-component of angular momentum of a system, is simply the sum of the z-components of angular momenta of all the parts of that system.

Secondly, any entity's total angular momentum  $J$  is unchanged by rotation.

Thirdly, the final state with maximum total spin is the linear sum of the original states with equal coefficients, equal in both magnitude and phase angle.

No other new principles of physics are introduced in this chapter. What follows is an algebraic tour de force on the summation and rotation of objects with extreme spin values.

Recall that a particle or system of particles can be characterized by  $j\hbar$ , its total angular momentum, and  $m\hbar$ , its component of angular momentum along any selected axis. Both  $j$  and  $m$  are quantized; they must be integer multiples of  $1/2$ , and  $m$  has  $(2j+1)$  possible values, ranging from  $-j$  to  $+j$  in integer steps. Absent external forces,  $j$  and  $m$  are conserved. Physicists generally select the z-axis, and only rarely mention the ubiquitous factor  $\hbar$ .

Each elementary particle has an intrinsic total spin  $j$  that never changes under any circumstances whatsoever. All elementary fermions have total spin  $1/2$ . Photons, gluons, and the  $Z^0$  and  $W^\pm$  bosons have spin 1. The Higgs boson has spin 0.

The spin of a system of several particles is the vector sum of the intrinsic spins of each particle plus any orbital angular momenta they might have. Orbital angular momenta are always integer multiples of  $\hbar$  (no half-integers here). While a system's total angular momentum  $J$  can be difficult to calculate, its total z-axis component  $M$  is simply the arithmetic sum of the  $m$  values of all its particles and of all their orbital motions.

We denote an angular momentum state by  $|j,m\rangle$ . The more common states are:

total spin 0:  $|0,0\rangle$

total spin  $1/2$ :  $|1/2,+1/2\rangle$ ;  $|1/2,-1/2\rangle$

total spin 1:  $|1,+1\rangle$ ;  $|1,0\rangle$ ;  $|1,-1\rangle$

Our objective here is to determine the projection of state  $|j,m\rangle$  onto all possible states  $|j^*,m^*\rangle$  in a

coordinate system whose axes are rotated by operator  $R$  relative to our original coordinate axes.

Rotating coordinate axes cannot change  $j$ , because angular momentum is conserved. But rotation does change the coordinate axes and the components of  $j$  along the new axes, which mixes  $m$ -states. The projection amplitudes we seek are the components of the rotation matrix:

$$\langle j, m^* | R | j, m \rangle$$

As discussed in Chapter 17, all possible rotations can be accomplished with a specific sequence of rotations by some set of three Euler angles. The rotation sequence, from coordinate system  $S$  to coordinate system  $T$ , is defined as:

1. Rotate about the  $z$ -axis by angle  $\beta$ .
2. Rotate about the new  $x$ -axis by angle  $\alpha$ .
3. Rotate about the new  $z$ -axis by angle  $\gamma$ .

(In various chapters, Feynman switches between  $x$  and  $y$  as the second Euler rotation axis. Either choice is valid, but for consistency, I always use the  $x$ -axis in step 2.)

Therefore, we can represent any rotation matrix  $R$  by:

$$R = R_z(\gamma) R_x(\alpha) R_z(\beta)$$

$R_z(\beta)$  is particularly simple:

$$R_z(\beta) |j, m\rangle = \exp\{im\beta\} |j, m\rangle$$

$$\langle j, m^* | R_z(\beta) | j, m \rangle = \exp\{im\beta\} \delta_{m, m^*}$$

Since  $R_z$  is a diagonal matrix, we can simplify the equation for the components of  $R$ .

$$\begin{aligned} R_{kn} &= \sum_{lj} R_{z,kl}(\gamma) R_{x,lj}(\alpha) R_{z,jn}(\beta) \\ R_{kn} &= \sum_l R_{z,kl}(\gamma) R_{x,ln}(\alpha) \exp\{in\beta\} \\ R_{kn} &= \exp\{ik\gamma\} R_{x,kn}(\alpha) \exp\{in\beta\} \end{aligned}$$

The only challenge then is calculating  $R_x(\alpha)$ .

In Chapter 17, we calculated spin 1 rotation matrices based on the assumption that a spin 1 system can be represented as the sum of two spin 1/2 components, whose rotation matrices we calculated in Chapter 8. We will now employ the same assumption to address even larger total spin values. Feynman offers no proof of this assumption. However, since the spins of all elementary particles are 1/2 and 1, any entity with a larger spin must actually be a physical combination of particles with lesser spins. Rotating the larger entity is the same physical act as rotating all of its parts.

Starting in V3p18-10, Feynman derives the rotation matrix for spin 3/2, and later derives it for any spin value.

## Rotation Matrix for Spin 3/2

A specific example of a spin 3/2 entity is deuterium, an atom comprised of three spin 1/2 particles: an electron orbiting a nucleus containing one proton and one neutron. By selecting three distinct particles, we avoid the complications of identical particle effects.

We represent each of the 8 possible spin states of deuterium by the m values of the three particles, with one being:  $|+->$ . For this derivation, it does not matter which particle has which spin; we are just dealing with three spins.

We are interested in the 8 states of definite total angular momentum  $j=3/2$ . These can be grouped in 4 distinct sets depending on their number of + spins:

$$\begin{aligned} |\psi_3>: m = +3/2: |+++> \\ |\psi_2>: m = +1/2: \{|++-> + |+-> + |-++>\}/\sqrt{3} \\ |\psi_1>: m = -1/2: \{|--> + |-+-> + |+-->\}/\sqrt{3} \\ |\psi_0>: m = -3/2: |---> \end{aligned}$$

In  $|\psi_2>$  and  $|\psi_1>$ , the three combinations are equally weighted. By symmetry, each particle's spin has the same effect. The factor of  $1/\sqrt{3}$  normalizes  $|\psi_2>$  and  $|\psi_1>$ .

It is very important to note that the three combinations are all added; none are subtracted. This is because the total spin is 3/2 only when each combination has the same sign. Let's understand why by considering a different combination:

$$|Q> = \{|++-> - |+-> + |-++>\}/\sqrt{3}$$

Now focus on the first two terms by defining state  $|V>$ .

$$|V> = |++-> - |+->$$

The spin of the first particle is the same in both terms, but the spins of the second and third particles reverse. Those two spins combine as:

$$|+-> - |->$$

This is a singlet state of spin 0, just as we saw in positronium and also in the hyperfine splitting of hydrogen. This means  $|V>$  is a combination of spin 0 and spin 1/2, which makes its total spin  $j=1/2$ . This makes  $|Q>$  a combination of  $j=1/2$  and  $j=3/2$ , which is not a pure  $j=3/2$  state. Since rotation cannot change  $j$ , a state like  $|Q>$  is not a possible outcome of rotating a spin 3/2 particle. The only possible outcomes have all combining terms with equal coefficients.

This is a general principle that applies to summing any spins: the final state with maximum total spin is the combination of the original spins with equal coefficients.

Next, let's confirm that the four  $|\psi\rangle$  states have the indicated m values. If each is rotated by angle  $\beta$  about the z-axis, the state is multiplied by  $\exp\{+i\beta/2\}$  for each particle — electron, proton, or neutron — with spin up (+1/2) and by  $\exp\{-i\beta/2\}$  for each particle with spin down (−1/2). The total impact of this rotation is:

$$\begin{aligned} R_z(\beta) |\psi_3\rangle &= \exp\{i\beta(+1+1+1)/2\} |\psi_3\rangle \\ R_z(\beta) |\psi_3\rangle &= \exp\{i\beta(+3/2)\} |\psi_3\rangle \\ \text{m of } |\psi_3\rangle &= +3/2 \end{aligned}$$

$$\begin{aligned} R_z(\beta) |\psi_2\rangle &= \exp\{i\beta(+1+1-1)/2\} |\psi_2\rangle \\ R_z(\beta) |\psi_2\rangle &= \exp\{i\beta(+1/2)\} |\psi_2\rangle \\ \text{m of } |\psi_2\rangle &= +1/2 \end{aligned}$$

$$\begin{aligned} R_z(\beta) |\psi_1\rangle &= \exp\{i\beta(+1-1-1)/2\} |\psi_1\rangle \\ R_z(\beta) |\psi_1\rangle &= \exp\{i\beta(-1/2)\} |\psi_1\rangle \\ \text{m of } |\psi_1\rangle &= -1/2 \end{aligned}$$

$$\begin{aligned} R_z(\beta) |\psi_0\rangle &= \exp\{i\beta(-1-1-1)/2\} |\psi_0\rangle \\ R_z(\beta) |\psi_0\rangle &= \exp\{i\beta(-3/2)\} |\psi_0\rangle \\ \text{m of } |\psi_0\rangle &= -3/2 \end{aligned}$$

Recall from Chapter 17 these amplitudes for Euler angle rotations of spin 1/2 particles:

$$\begin{aligned} a &= \langle +T|S \rangle = \cos(\alpha/2) \exp\{+i(\beta+\gamma)/2\} \\ b &= \langle -T|S \rangle = i \sin(\alpha/2) \exp\{+i(\beta-\gamma)/2\} \\ c &= \langle +T|-S \rangle = i \sin(\alpha/2) \exp\{-i(\beta-\gamma)/2\} \\ d &= \langle -T|-S \rangle = \cos(\alpha/2) \exp\{-i(\beta+\gamma)/2\} \end{aligned}$$

We next calculate four key components of rotation matrix R by assuming each of the three particles is rotated independently.

$$\langle \psi_3 | R | \psi_3 \rangle = \langle +++ | R | +++ \rangle = \mathbf{a \cdot a \cdot a} = \mathbf{a^3}$$

$$\begin{aligned} \langle \psi_3 | R | \psi_2 \rangle &= \langle +++ | R \{ |+-> + |-+> + |-++> \} / \sqrt{3} \\ &= \{aac + aca + caa\} / \sqrt{3} = \mathbf{a^2 c \sqrt{3}} \end{aligned}$$

$$\begin{aligned} \langle \psi_2 | R | \psi_3 \rangle &= \{ \langle ++- | + \langle +-+ | + \langle -++ | \} | R | +++ \rangle / \sqrt{3} \\ &= \{aab + aba + baa\} / \sqrt{3} = \mathbf{a^2 b \sqrt{3}} \end{aligned}$$

$$\begin{aligned} \langle \psi_2 | R | \psi_2 \rangle &= \{ \langle ++- | + \langle +-+ | + \langle -++ | \} R | +-> / 3 \\ &\quad + \{ \langle ++- | + \langle +-+ | + \langle -++ | \} R | +-> / 3 \\ &\quad + \{ \langle ++- | + \langle +-+ | + \langle -++ | \} R | -++> / 3 \end{aligned}$$

$$\begin{aligned}
&= \{aad + abc + bac\}/3 \\
&\quad + \{acb + ada + bca\}/3 \\
&\quad + \{cab + cba + daa\}/3
\end{aligned}$$

$$= \mathbf{a^2d + 2abc}$$

There are 12 more components to go. But I have a simple way to obtain these using symmetry.

Note that swapping + and – spins (denoted: + <=> –) on the right (the Dirac bra) exchanges:

$$|\psi_3\rangle \Leftrightarrow |\psi_0\rangle$$

$$|\psi_2\rangle \Leftrightarrow |\psi_1\rangle$$

$$a \Leftrightarrow c$$

$$b \Leftrightarrow d$$

This gives us four more components.

$$\langle\psi_3|R|\psi_3\rangle = a^3 \Leftrightarrow \langle\psi_3|R|\psi_0\rangle = c^3$$

$$\langle\psi_3|R|\psi_2\rangle = a^2c\sqrt{3} \Leftrightarrow \langle\psi_3|R|\psi_1\rangle = c^2a\sqrt{3}$$

$$\langle\psi_2|R|\psi_3\rangle = a^2b\sqrt{3} \Leftrightarrow \langle\psi_2|R|\psi_0\rangle = c^2d\sqrt{3}$$

$$\langle\psi_2|R|\psi_2\rangle = a^2d + 2abc \Leftrightarrow$$

$$\langle\psi_2|R|\psi_1\rangle = c^2b + 2cda$$

Now, note that swapping + and – spins on the left (the Dirac ket) exchanges:

$$\langle\psi_3| \Leftrightarrow \langle\psi_0|$$

$$\langle\psi_2| \Leftrightarrow \langle\psi_1|$$

$$a \Leftrightarrow b$$

$$c \Leftrightarrow d$$

This gives us the last eight components.

$$\langle\psi_2|R|\psi_3\rangle = a^2b\sqrt{3} \Leftrightarrow \langle\psi_1|R|\psi_3\rangle = b^2a\sqrt{3}$$

$$\langle\psi_2|R|\psi_2\rangle = a^2d + 2abc \Leftrightarrow \langle\psi_1|R|\psi_2\rangle = b^2c + 2bad$$

$$\langle\psi_2|R|\psi_1\rangle = c^2b + 2cda \Leftrightarrow \langle\psi_1|R|\psi_1\rangle = d^2a + 2dcb$$

$$\langle\psi_2|R|\psi_0\rangle = c^2d\sqrt{3} \Leftrightarrow \langle\psi_1|R|\psi_0\rangle = d^2c\sqrt{3}$$

$$\langle\psi_3|R|\psi_3\rangle = a^3 \Leftrightarrow \langle\psi_0|R|\psi_3\rangle = b^3$$

$$\langle\psi_3|R|\psi_2\rangle = a^2c\sqrt{3} \Leftrightarrow \langle\psi_0|R|\psi_2\rangle = b^2d\sqrt{3}$$

$$\langle\psi_3|R|\psi_1\rangle = c^2a\sqrt{3} \Leftrightarrow \langle\psi_0|R|\psi_1\rangle = d^2b\sqrt{3}$$

$$\langle\psi_3|R|\psi_0\rangle = c^3 \Leftrightarrow \langle\psi_0|R|\psi_0\rangle = d^3$$

All these results are tabulated in matrix form at the end of this chapter.

## Composition of |J,M> Sums

For clarity, we will now use J and M for the quantum numbers of a combined system, and j and m for the quantum numbers of its constituent particles.

Summing angular momenta from several sources entails a great deal of algebra and combinatorics, and the three key principles given at the start of this chapter. These are:

Firstly,  $M = \sum_n m_n$ . M, the z-component of angular momentum of a system, is simply the sum of the z-components of angular momenta of all the parts of that system.

Secondly, any entity's total angular momentum J is unchanged by rotation.

Thirdly, the final state with maximum total spin is the linear sum of the original states with equal coefficients, equal in both magnitude and phase angle.

We found that the final state's total spin is reduced when the spin combinations are summed with opposite polarities. In particular, one possible sum of two spins with  $j=1/2$  is:

$$|+-\rangle - |-+\rangle$$

Clearly this sum has  $m=0$ , because each term has two opposite spins. Less clear, but no less true, is the fact that this is a singlet state with  $j=0$ , due to the minus sign between the two terms. If we add the two terms instead of subtracting them,  $m$  would still equal 0 but  $j$  would now equal 1. Adding the two terms results in the  $m=0$  part of a spin 1 triplet.

The same idea pertains to more complex situations. If we combine five spin  $1/2$  particles, the possible  $|J, M\rangle$  outcomes are:

$$J = 5/2; M = +5/2, +3/2, +1/2, -1/2, -3/2, -5/2$$

$$J = 3/2; M = +3/2, +1/2, -1/2, -3/2$$

$$J = 1/2; M = +1/2, -1/2$$

We see that  $M=+1/2$  occurs for three different J values. For  $M=+1/2$  and  $J=5/2$ , all terms with all different combinations of three spin up and two spin down are combined with the same polarity. When one term has the opposite polarity, a pair of particles form a spin 0 singlet, leaving the other three particles summing with the same polarity to form a  $J=3/2$  state. In other combinations of mixed polarities, two pairs of particles form two spin-zero singlet states, leaving the last particle to supply all of the angular momentum, resulting in a  $J=1/2$  state.

We can write the following general equation describing how  $n$  states with  $|j_n, m_n\rangle$  combine to form  $|J, M\rangle$  states:

$$|J, M\rangle = \sum_n \langle J, M | j_n, m_n \rangle |j_n, m_n\rangle$$

with the constraint that  $M = \sum_n m_n$ .

The numbers  $\langle J, M | j_n, m_n \rangle$  are called *Clebsch-Gordon coefficients*. Tables of these coefficients are available online.

## Rotation of Any Spin

In V3p18-19, Feynman derives the matrix for a rotation about the y-axis by angle  $\theta$  for any spin value  $j$ . That rotation corresponds to an Euler rotation with angles:  $\alpha=\theta$ ,  $\beta=+\pi/2$ , and  $\gamma=-\pi/2$ . For spin  $1/2$ , the rotation tables yield:

$$\begin{aligned} a &= \langle +T|+S \rangle = +\cos(\alpha/2) \\ b &= \langle -T|+S \rangle = -\sin(\alpha/2) \\ c &= \langle +T|-S \rangle = +\sin(\alpha/2) \\ d &= \langle -T|-S \rangle = +\cos(\alpha/2) \end{aligned}$$

In V3p18-12, Feynman quotes this result:

$$\begin{aligned} \langle j, m^* | R_y(\theta) | j, m \rangle \\ = [(j+m)! (j-m)! (j+m^*)! (j-m^*)!]^{1/2} \\ \times \sum_k \{ (-1)^k a^{(2j+m^*-m-2k)} c^{(m-m^*+2k)} / D \} \end{aligned}$$

$$\text{where } D = [(m-m^*+k)! (j+m^*-k)! (j-m-k)! k!]$$

The above sum extends over all values of  $k$  for which all factorials have arguments  $\geq 0$ . Note that  $0!=1$ , by definition.

This is the ugliest equation I have ever seen, and that is saying a lot.

I provide here a few tables for the simpler cases.

- For **two non-identical spin 1/2** particles:

$$\begin{aligned} J=1, M=+1: & | +1/2, +1/2 \rangle \\ J=1, M=0: & \{ | +1/2, -1/2 \rangle + | -1/2, +1/2 \rangle \} / \sqrt{2} \\ J=1, M=-1: & | -1/2, -1/2 \rangle \\ J=0, M=0: & \{ | +1/2, -1/2 \rangle - | -1/2, +1/2 \rangle \} / \sqrt{2} \end{aligned}$$

- For a **spin 1/2** and a **spin 1** particle:

$$\begin{aligned} J=3/2, M=+3/2: & | +1/2, +1 \rangle \\ J=3/2, M=+1/2: & | +1/2, 0 \rangle \sqrt{2/3} + | -1/2, +1 \rangle / \sqrt{3} \\ J=3/2, M=-1/2: & | -1/2, 0 \rangle \sqrt{2/3} + | +1/2, -1 \rangle / \sqrt{3} \\ J=3/2, M=-3/2: & | -1/2, -1 \rangle \\ J=1/2, M=+1/2: & | +1/2, 0 \rangle \sqrt{3} - | -1/2, +1 \rangle / \sqrt{2/3} \\ J=1/2, M=-1/2: & | +1/2, -1 \rangle \sqrt{2/3} - | -1/2, 0 \rangle / \sqrt{3} \end{aligned}$$

- For **two non-identical spin 1** particles:

$$\begin{aligned}
J=2, M=+2: & | +1, +1 \rangle \\
J=2, M=+1: & \{ | +1, 0 \rangle + | 0, +1 \rangle \} / \sqrt{2} \\
J=2, M=0: & \{ | +1, -1 \rangle + | -1, +1 \rangle + 2 | 0, 0 \rangle \} / \sqrt{6} \\
J=2, M=-1: & \{ | 0, -1 \rangle + | -1, 0 \rangle \} / \sqrt{2} \\
J=2, M=-2: & | -1, -1 \rangle \\
J=1, M=+1: & \{ | +1, 0 \rangle - | 0, +1 \rangle \} / \sqrt{2} \\
J=1, M=0: & \{ | +1, -1 \rangle - | -1, +1 \rangle \} / \sqrt{2} \\
J=1, M=-1: & \{ | 0, -1 \rangle - | -1, 0 \rangle \} / \sqrt{2} \\
J=0, M=0: & \{ | +1, -1 \rangle + | -1, +1 \rangle - | 0, 0 \rangle \} / \sqrt{3}
\end{aligned}$$

## Spin 3/2 Rotation Matrix

Below is the complete rotation matrix for spin 3/2 particles for any set of Euler angles:

|             |                                  |                                  |                                  |                                  |
|-------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
| <b>m</b>    | <b>m in S Frame</b>              |                                  |                                  |                                  |
| <b>in T</b> | <b>+3/2</b>                      | <b>+1/2</b>                      | <b>-1/2</b>                      | <b>-3/2</b>                      |
| <b>+3/2</b> | <b><math>a^3</math></b>          | <b><math>a^2c\sqrt{3}</math></b> | <b><math>ac^2\sqrt{3}</math></b> | <b><math>c^3</math></b>          |
| <b>+1/2</b> | <b><math>a^2b\sqrt{3}</math></b> | <b><math>a^2d+2abc</math></b>    | <b><math>c^2b+dac</math></b>     | <b><math>c^2d\sqrt{3}</math></b> |
| <b>-1/2</b> | <b><math>b^2a\sqrt{3}</math></b> | <b><math>b^2c+2bad</math></b>    | <b><math>d^2a+cbd</math></b>     | <b><math>d^2c\sqrt{3}</math></b> |
| <b>-3/2</b> | <b><math>b^3</math></b>          | <b><math>b^2d\sqrt{3}</math></b> | <b><math>bd^2\sqrt{3}</math></b> | <b><math>d^3</math></b>          |

Here we use these amplitudes for Euler angle rotations of spin 1/2 particles:

$$\begin{aligned}
a &= \langle +T | +S \rangle = \cos(\alpha/2) \exp\{+i(\beta+\gamma)/2\} \\
b &= \langle -T | +S \rangle = i \sin(\alpha/2) \exp\{+i(\beta-\gamma)/2\} \\
c &= \langle +T | -S \rangle = i \sin(\alpha/2) \exp\{-i(\beta-\gamma)/2\} \\
d &= \langle -T | -S \rangle = \cos(\alpha/2) \exp\{-i(\beta+\gamma)/2\}
\end{aligned}$$

## Chapter 28 Review: Key Ideas

Summing angular momenta from several sources entails a great deal of algebra and combinatorics, and three key principles that are:

Firstly,  $M = \sum_n m_n$ .  $M$ , the z-component of angular momentum of a system, is simply the sum of the z-components of angular momenta of all the parts of that system.

Secondly, any entity's total angular momentum  $J$  is unchanged by rotation.

Thirdly, the final state with maximum total spin is the linear sum of the original states with equal coefficients, equal in both magnitude and phase angle.

# Chapter 29

## Electron Wave Functions in Hydrogen

---

In V3p19-1, Feynman begins this lecture by saying:

“The most dramatic success in the history of the quantum mechanics was the understanding of the details of the spectra of some simple atoms and the understanding of the periodicities which are found in the [Periodic Table] of chemical elements.”

This chapter derives the atomic structure of hydrogen, including all the excited states of its electron. The next chapter qualitatively explains the structure of the Periodic Table.

We will make some simplifying assumptions. We will do our analysis in the proton’s center of mass frame — that is, we shall assume the proton is stationary. Also, we will use non-relativistic equations and assume that all particle spins remain fixed.

These are excellent approximations, in this case, that deviate only modestly from a completely precise, but much more complex, analysis.

Angular momentum conservation requires the sum of (spin angular momentum) plus (orbital angular momentum) to remain constant. Since we assume all particle spins are constant, this means the electron’s orbital angular momentum must also be constant.

### Schrödinger Equation for Hydrogen

With these approximations, let wave function  $\psi(x,y,z,t)$  be the amplitude to find the electron of a hydrogen atom at position  $(x,y,z)$  at time  $t$ . The evolution of  $\psi$  through space and time is given by Schrodinger’s equation (see Chapter 22):

$$i\hbar d\psi/dt = -\hbar^2/2\mu \nabla^2\psi + V \psi$$

Here  $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ , and  $V$  is the electrostatic potential due to the proton. We will use  $m$  to denote the magnetic quantum number ( $z$ -component of orbital angular momentum), and  $\mu$  to denote the electron mass.

The zero-point of any form of potential energy is arbitrary, because only potential differences have physical consequences. We make the standard choice:  $V=0$  for an electron that is infinitely far from the proton. We also set the center of our coordinate system at the stationary proton. The potential is

then:

$$V(r) = -e^2/r$$

where  $r$  is the electron's radial coordinate, and  $e^2 = q_e^2 / 4\pi\epsilon_0$ .

Our equation is now:

$$\hbar \frac{d\psi}{dt} = -\frac{\hbar^2}{2\mu} \nabla^2 \psi - (e^2/r) \psi$$

We seek stationary solutions of this equation, which means solutions of the form:

$$\psi(x,y,z,t) = \psi(x,y,z) \exp\{-iEt/\hbar\}$$

Plugging that into our equation results in:

$$E \psi = -\frac{\hbar^2}{2\mu} \nabla^2 \psi - (e^2/r) \psi$$

Since  $V$  is a function of radius, polar coordinates  $(r, \theta, \phi)$  will be more convenient here than rectilinear coordinates  $(x, y, z)$ . These coordinate systems are related by:

$$x = r \sin\theta \cos\phi$$

$$y = r \sin\theta \sin\phi$$

$$z = r \cos\theta$$

Translating  $\nabla^2$  into polar coordinates results in a rather messy equation. (However, in rectilinear coordinates, there are three interconnected equations that are much harder to solve.)

## Case of Spherical Symmetry

Let's begin with the simplest case: a spherically symmetric electron wave function, which is called an *s-orbit*. In this case,  $\psi$  is a function of radius  $r$ , but  $\psi$  does not vary with either the polar or azimuthal angles. This means wave function  $\psi$  will be a state of zero orbital angular momentum, since as we are well aware by now, states of non-zero angular momentum do change under rotation.

Spherical symmetry simplifies  $\nabla^2$ . For any spherically symmetric function  $f$ :

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

$$\nabla^2 f = \frac{\partial^2 (rf)}{\partial r^2} / r$$

Our equation becomes:

$$E \psi = -\frac{\hbar^2}{2\mu} \left\{ \frac{\partial^2 (r\psi)}{\partial r^2} \right\} / r - (e^2/r) \psi$$

$$\frac{\partial^2 (r\psi)}{\partial r^2} / r = - (2\mu/\hbar^2) (E + e^2/r) \psi$$

In V3p19-3, Feynman makes two substitutions to reduce clutter. These substitutions amount to

analyzing the physics in natural units — measuring quantities in units that are characteristic of the natural phenomenon, rather than anthropic units, such as meters and seconds. The substitutions are:

$$\rho = r / r_b, \text{ which means: } r = r_b \rho$$

$$\varepsilon = E / R_y, \text{ which means: } E = R_y \varepsilon$$

where:

$$r_b = \hbar^2 / \mu e^2 = \text{Bohr radius} = 0.529 \text{ angstroms}$$

$$R_y = \mu e^4 / 2 \hbar^2 = 1 \text{ Rydberg} = 13.6 \text{ eV}$$

Note that:

$$r_b R_y = [\hbar^2 / \mu e^2] [\mu e^4 / 2 \hbar^2] = e^2 / 2$$

In natural units, our equation is:

$$\partial^2(\rho\psi) / \partial \rho^2 / (\rho r_b^2) = -(2\mu / \hbar^2) \{ R_y \varepsilon + e^2 / \rho r_b \} \psi$$

$$\begin{aligned} \partial^2(\rho\psi) / \partial \rho^2 / \rho &= -r_b (2\mu / \hbar^2) \{ r_b R_y \varepsilon + e^2 / \rho \} \psi \\ &= -(\hbar^2 / \mu e^2) (2\mu / \hbar^2) \{ r_b R_y \varepsilon + e^2 / \rho \} \psi \\ &= -(2/e^2) \{ [e^2/2] \varepsilon + e^2 / \rho \} \psi \end{aligned}$$

$$\partial^2(\rho\psi) / \partial \rho^2 / \rho = -(\varepsilon + 2/\rho) \psi$$

$$\partial^2(\rho\psi) / \partial \rho^2 = -(\varepsilon + 2/\rho) \rho\psi$$

How do we solve this differential equation for  $\psi$ ?

Unfortunately, there is no general procedure for solving differential equations. This is because there is no general procedure for solving an arbitrary integral. But, over the years, people have developed a lot of tricks. The more tricks you learn and the more experience you gain, the more successful you become at this challenge. This is one reason to learn from the best.

Recall from Chapter 22 that, at large distances  $r$  from the proton, the electron's wave function  $\psi$  decreases exponentially with  $r$ . This is because the electron is bound and  $V > E$  at large  $r$ . We therefore factor out the exponential by making this substitution:

$$\rho\psi = g(\rho) \exp\{-\beta\rho\}$$

Let's calculate the second partial derivative of the right hand side of the above equation.

$$\begin{aligned} \partial( g(\rho) \exp\{-\beta\rho\} ) / \partial \rho \\ = \exp\{-\beta\rho\} \partial g / \partial \rho - \beta g \exp\{-\beta\rho\} \end{aligned}$$

$$\begin{aligned} \partial^2( g(\rho) \exp\{-\beta\rho\} ) / \partial \rho^2 \\ = -\beta \exp\{-\beta\rho\} \partial g / \partial \rho + \exp\{-\beta\rho\} \partial^2 g / \partial \rho^2 \end{aligned}$$

$$-\beta \partial g / \partial \rho \exp \{-\beta \rho\} + \beta^2 g \exp \{-\beta \rho\}$$

$$= [-2\beta \partial g / \partial \rho + \partial^2 g / \partial \rho^2 + \beta^2 g] \exp \{-\beta \rho\}$$

Our differential equation becomes:

$$\partial^2 (g \exp \{-\beta \rho\}) / \partial \rho^2 = -(\epsilon + 2/\rho) g \exp \{-\beta \rho\}$$

$$[-2\beta \partial g / \partial \rho + \partial^2 g / \partial \rho^2 + \beta^2 g] = -(\epsilon + 2/\rho) g$$

$$\partial^2 g / \partial \rho^2 - 2\beta \partial g / \partial \rho + (\beta^2 + \epsilon + 2/\rho) g = 0$$

Believe it or not, we really are making progress.

Now, since  $\beta$  is an arbitrary parameter in our substitution, we can set it to any value we wish. Let's choose  $\beta^2 = -\epsilon$ , reducing our equation to:

$$\partial^2 g / \partial \rho^2 - 2\beta \partial g / \partial \rho + 2g/\rho = 0$$

If not for the  $\rho$  in the denominator of the third term, this would be a linear differential equation with constant coefficients that we learned how to solve in Chapters 12 through 14 of *Feynman Simplified 1B*. Here we have a more complex differential equation, but it can be solved with a power series. Let:

$$g(\rho) = \sum_k a_k \rho^k$$

where the  $a$ 's are constants, and the sum might extend from  $k=0$  to  $k=+\infty$ . The above sum is called a *Taylor series expansion*. Almost any physically realistic function can be represented by a Taylor series, which makes this a standard technique in every theoretical physicist's toolbox. This technique works well when the coefficients  $a_k$  approach zero for large  $k$ .

Putting the Taylor series into our equation yields:

$$0 = \sum_k a_k \{k(k-1)\rho^{k-2} - 2\beta k\rho^{k-1} + 2\rho^{k-1}\}$$

Note that for  $k=0$ , we get  $0=2a_0/\rho$ , which means the Taylor series that solves this differential equation must have  $a_0=0$  and we can start the summation at  $k=1$ .

We have above three independent sums, allowing us to redefine the summation index of each term to obtain the following:

$$0 = \sum_n \{(n+1)na_{n+1} - 2\beta na_n + 2a_n\} \rho^{n-1}$$

Here the sum is from  $n=1$  to  $n=+\infty$ . In the second and third terms, we set  $n=k$ . In the first term we set  $n=k-1$ , and note that  $k=1$  ( $n=0$ ) makes no contribution to the sum. This allows all three summations to begin at  $n=1$ .

The above equation is valid for all values of  $\rho$ . This can only be true if the coefficient of each power of  $\rho$  is zero. This is an important rule for polynomials that is well worth remembering. The proof of this rule is given at the end of this chapter. This means:

$$0 = \{(n+1)a_{n+1} - 2\beta a_n + 2a_n\}, \text{ for all } n$$

Rearranging yields:

$$a_{n+1} = a_n \{2(\beta n - 1)/n(n+1)\}$$

This equation is valid for all  $n \geq 1$ . With any choice of  $a_1$ , we can recursively calculate  $a_2$ , then  $a_3$ , etc. We can calculate all the  $a_n$ 's in terms of the variable  $\beta$ . Recall that  $\beta$  is related to the electron's energy by:

$$\beta = \sqrt{-\epsilon} = \sqrt{-E/R_y}$$

$E$  must be negative for a bound electron. Hence, any value of  $E < 0$  yields a real value of  $\beta$ , a valid Taylor series for  $g(\rho)$ , and a mathematical solution to Schrödinger's equation of the form:

$$\psi(\rho) = g(\rho) \exp\{-\beta\rho\} / \rho$$

In V3p19-4, Feynman says: "We have a solution, but what does it represent physically?"

Consider how this solution behaves at large  $\rho$ , where we know that any real wave function must decrease exponentially. For large  $\rho$ , the dominant terms in a polynomial generally come from the highest powers of  $\rho$ . For these, we can approximate the equation for the  $a_n$ 's by:

$$a_{n+1} = a_n \{2\beta/(n+1)\}$$

Each higher term is multiplied by  $2\beta$  and divided by its term number. For large  $n$  this means:

$$a_{n+1} = a_1 (2\beta)^{n+1} / 2\beta(n+1)!$$

$$g(\rho) \sim \sum_n [(2\beta)^n/n!] \rho^n$$

But this is just the power series expansion of an exponential. Recall that:

$$\exp\{x\} = \sum_n x^n / n!$$

Thus for large  $\rho$ , the solution we found for  $g(\rho)$  increases exponentially as:

$$g(\rho) \sim \exp\{2\beta\rho\}$$

Putting this into the equation for  $\psi$  yields, for large  $\rho$ :

$$\begin{aligned} \rho\psi &= g(\rho) \exp\{-\beta\rho\} \\ \psi(\rho) &\sim \exp\{+\beta\rho\} / \rho \end{aligned}$$

This is impossible. No physically realistic wave function can increase exponentially as its independent variable grows without limit.

While perhaps initially distressing, this derivation confirms that only certain electron energies can lead to solutions of Schrödinger's equation. To obtain a bound solution, the Taylor series must be truncated in order to avoid its divergence at large  $n$ . This means each solution must have:

$$a_{n+1} = a_n \{2(\beta n - 1)/n(n+1)\} = 0, \text{ for some } n$$

If  $a_{n+1} = 0$ , all  $a$ 's with larger indices will also be zero, thereby preventing the wave function from exploding exponentially at large radii.

Therefore, for some  $n$ :

$$\begin{aligned} \beta n &= 1 \\ \sqrt{(-E/R_y)} &= 1/n \\ E &= -R_y/n^2 \end{aligned}$$

Physically realistic electron wave functions correspond to integer values of  $n$ , which is called the *principal quantum number*. Each value of  $n$  corresponds to a bound energy ( $E < 0$ ) given by the above equation.

Recall that this is the same equation we derived in *Feynman Simplified 3A* Chapter 3.

There, we derived this equation assuming the circumference of an electron's orbit must be an integral number of wavelengths, which is equivalent to assuming the electron's orbital angular momentum must be an integral multiple of  $\hbar$ .

Here, we have derived this equation based on Schrödinger's equation of motion.

The complete spherically symmetric solutions are:

$$E_n = -R_y/n^2$$

$$\psi_n(\rho) = g_n(\rho) \exp\{-\rho/n\} / \rho$$

$$g_n(\rho) = \sum_k a_k \rho^k, \text{ sum over } k=1 \text{ to } k=n$$

$$a_{k+1} = -a_k \{2(1-k/n)/k(k+1)\}$$

Let's examine the solutions for a few values of  $n$ .

The minimum energy level, the ground state, has  $n=1$ , with:

$$\begin{aligned} E_1 &= -R_y \\ \psi_1(\rho) &= a_1 \exp\{-\rho\} \end{aligned}$$

The wave function  $\psi_1(\rho)$  peaks at the atom's center, at the proton, and decreases exponentially with radius. Recall that  $\rho$  is in units of the Bohr radius, so  $\psi_1(\rho)$  drops by a factor of  $e$  from the center to one Bohr radius.

In V3p19-5, Feynman says something that might be misleading. He says the electron “is most likely to be found right at the proton.” While no other single point has a higher amplitude than the center, it is important to recognize that  $\psi$  is an *amplitude density*. The magnitude of the amplitude squared must be integrated over a non-zero volume to obtain a meaningful probability. The volume of a shell of radius  $\rho$  and thickness  $d\rho$  equals  $4\pi\rho^2 d\rho$ . Hence, for  $\psi_1$ , the probability of the electron being a distance  $\rho$  from the atom's center is proportional to  $\rho^2 \exp\{-2\rho\}$ . The probability of finding the electron near  $\rho=0$  is zero. The electron's most likely radius is where the derivative of its probability distribution is zero.

$$\begin{aligned} 0 &= \partial(\rho^2 \exp\{-2\rho\})/\partial\rho \\ 0 &= 2\rho \exp\{-2\rho\} - 2\rho^2 \exp\{-2\rho\} \\ 0 &= 1 - \rho \\ \rho &= 1 \text{ Bohr radius} \end{aligned}$$

We can also check the average potential energy  $\langle V \rangle$  of this orbit. First, we must normalize  $\psi$ , by integrating  $\psi^*\psi$  multiplied by the volume element  $4\pi\rho^2 d\rho$  from  $\rho=0$  to  $\rho=+\infty$ . In integral tables, we find this definite integral:

$$\int x^n \exp\{-ax\} dx = n!/a^{n+1}; \text{ for } x \text{ from } 0 \text{ to } +\infty$$

$$\begin{aligned} 1 &= A \int \psi^* \psi 4\pi\rho^2 d\rho \\ 1 &= 4\pi A \int \exp\{-2\rho\} \rho^2 d\rho \\ 1 &= 4\pi A [2/2^3] \\ A &= 1/\pi \end{aligned}$$

Next, we integrate the normalized  $\psi^*\psi$  times the volume element times the potential over all  $\rho$ :

$$\begin{aligned} \langle V \rangle &= 4\pi A \int \exp\{-2\rho\} \rho^2 (-e^2/\rho r_b) d\rho \\ \langle V \rangle &= -(4e^2/r_b) \int \exp\{-2\rho\} \rho d\rho \\ \langle V \rangle &= -(4e^2/r_b) [1/4] \\ \langle V \rangle &= -e^2/r_b = -27.22 \text{ eV} = 2R_y \end{aligned}$$

As we discussed in ***Feynman Simplified 3A*** Chapter 3, the electron's kinetic energy equals  $-R_y$  ( $>0$ ), exactly  $-1/2$  times the average potential energy. The electron's total energy equals  $2R_y - 1R_y = 1R_y$  ( $<0$ ). The Virial theorem explains why a system at equilibrium in a  $1/r$  potential must have this ratio of potential to kinetic energy.

The next higher energy level has  $n=2$ . Here we have:

$$\begin{aligned} E_2 &= -R_y/4 = -3.4 \text{ eV} \\ a_2 &= -a_1 \{2(1-1/2)/2\} = -a_1/2 \\ \psi_2(\rho) &= a_1 (1 - \rho/2) \exp\{-\rho/2\} \end{aligned}$$

For  $n=3$ , we have:

$$E_3 = -R_y/9 = -1.5 \text{ eV}$$

$$a_2 = -a_1 \{2(1-1/3)/2\} = -2a_1/3$$

$$a_3 = -a_2 \{2(1-2/3)/6\} = -a_2/9 = +2a_1/27$$

$$\psi_3(\rho) = a_1 (1 - 2\rho/3 + 2\rho^2/27) \exp\{-\rho/3\}$$

These three s-orbits are illustrated in Figure 29-1, where the intensity of black is proportional to the electron's probability density.

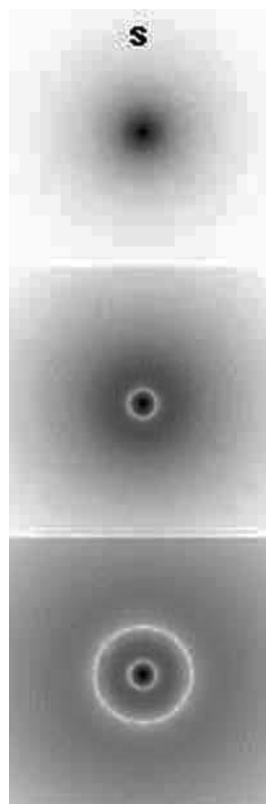


Figure 29-1 Hydrogen s-Orbits  
Top to Bottom:  $n=1, 2, 3$

We see in Figure 29-1 that the probability density of the  $n=1$  orbit drops continuously from the center outward. For  $n=2$ , the probability density passes through zero once (one white ring), while for  $n=3$ , it passes through zero twice.

## Orbits With Angular Momentum

Only s-orbits, which are spherically symmetric, have zero orbital angular momentum. Let's now examine orbits with non-zero orbital angular momentum.

As before,  $j$  is the total orbital angular momentum quantum number, and its component along any selected axis is the magnetic quantum number  $m$ . The total orbital angular momentum equals  $j\hbar$ , and its component equals  $m\hbar$ .

For principal quantum number  $n$ , the allowed values of  $j$  and  $m$  are:

$$j = 0, 1, 2 \dots (n-1)$$

$$m = -j, -(j-1) \dots 0 \dots +(j-1), +j$$

The states of various total orbital angular momenta are historically denoted by a single letter, as follows:

$$\begin{aligned} j=0: & s \\ j=1: & p \\ j=2: & d \\ j=3: & f \\ j=4: & g \\ j=5: & h \end{aligned}$$

There are no known elements that normally have electrons in orbits with  $j>3$ , although excited electrons may briefly occupy higher  $j$  orbits.

Now, we will solve the Schrödinger equation for wave functions that are not spherically symmetric. Recall this equation from earlier in this chapter.

$$\hbar \frac{d\psi}{dt} = -\frac{\hbar^2}{2\mu} \nabla^2 \psi - \left(\frac{e^2}{r}\right) \psi$$

where  $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$  in rectilinear coordinates.

For the spherically symmetric case, in polar coordinates, we used:

$$\nabla^2 f = \partial^2(rf)/\partial r^2$$

Now, for wave functions that vary with polar angle  $\theta$  and azimuthal angle  $\phi$ , we need the full polar coordinate form of  $\nabla^2$ , which is:

$$\begin{aligned} \nabla^2 f = & \partial^2(rf)/\partial r^2 \\ & + \left\{ \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) \frac{\partial}{\partial \theta} + \frac{(\partial^2 f / \partial \phi^2) / \sin \theta}{r^2 \sin \theta} \right\} \end{aligned}$$

Many courses in quantum mechanics tackle this formidable equation head-on, but Feynman has a better way. We will use what we have learned about angular momentum.

Firstly, since the electrostatic potential is spherically symmetric (varying only as a function of  $r$ ), it cannot exert a torque on the electron (the force is radial and has zero lever arm). This means total angular momentum is conserved about every axis. Since we are assuming that the electron and proton spins remain fixed, this means total orbital angular momentum is conserved about every axis — the electron's  $j$  quantum number is constant in time and invariant under rotation.

Now consider an electron that is somewhere along the  $+z$ -axis at a distance  $r>0$  from the center. The key insight is: this electron's magnetic quantum number  $m$ , its  $z$ -component of orbital angular momentum, must be zero. We can see why this is true by recalling the classical equation for angular momentum:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

$$L_z = x p_y - y p_x$$

On the z-axis,  $x=y=0$ , the momentum has no lever arm on the z-axis, and therefore  $L_z=0$ .

Define  $F_j(r)$  to be the amplitude for an electron to be on the z-axis, at distance  $r$  from the center, with total orbital angular momentum  $j$ , and  $m=0$  along the z-axis.

By knowing the amplitude in one direction, we can calculate the amplitude in any direction using rotation operators. Define the  $z^*$ -axis to be the direction of the original z-axis after the operation  $R_y(\theta)R_z(\phi)$ , as illustrated in Figure 29-2.

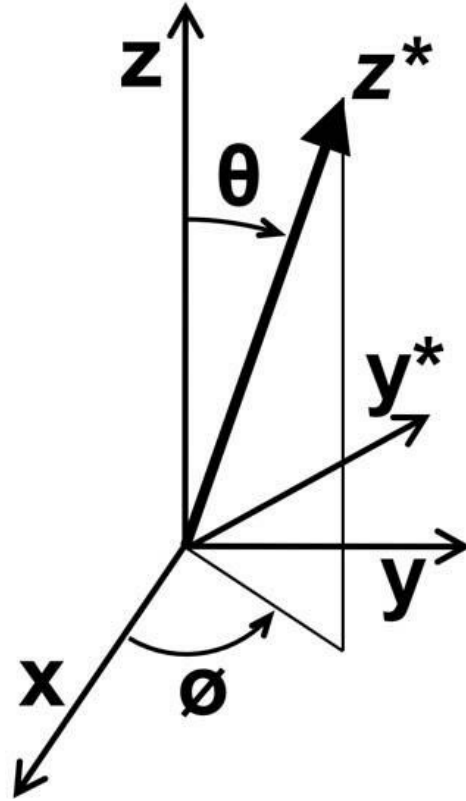


Figure 29-2.  $z^* = R_y(\theta)R_z(\phi) z$

As shown in the figure,  $R_y(\theta)R_z(\phi)$  begins with a rotation by angle  $\phi$  about the z-axis, followed by a rotation by angle  $\theta$  about the new  $y^*$ -axis. We know that rotations do not change  $r$  and  $j$ , but they do mix  $m$ -states. Since  $r$  and  $j$  are invariant under rotation, so is  $F_j(r)$ .

The amplitude for an electron to be on the z-axis, at distance  $r$  from the center, in orbital angular momentum state  $|j,m\rangle$  equals the product of {projection amplitude of  $|j,m\rangle$  along  $z$  onto  $|j,0\rangle$  along  $z^*$ } multiplied by  $\{F_j(r)$ , amplitude to be at distance  $r$  along  $z^*$ -axis with  $|j,0\rangle\}$ . This is written:

$$\psi_{j,m}(r) = Y_{j,m}(\theta,\phi) F_j(r)$$

$$Y_{j,m}(\theta,\phi) = \langle j,0 | R_y(\theta)R_z(\phi) | j,m \rangle$$

The electron states given by  $\psi_{j,m}(r)$  should be considered basis states: each  $\psi_{j,m}(r)$  is a solution of the Schrödinger equation, and so is any linear combination with the same  $j$ .

This particular combination of angular functions, called *spherical harmonics*, occurs frequently in the analysis of wave phenomena using polar coordinates. Since both atomic scale and cosmic scale entities have spherical geometries, the spherical harmonics  $Y_{j,m}(\theta,\phi)$  are used in both quantum mechanics and in classical physics.

Another common representation of spherical harmonics uses the *associated Legendre functions*  $P_j^m(\cos\theta)$  :

$$P_j^m(\cos\theta) \exp\{im\phi\} = Y_{j,m}(\theta,\phi)$$

It is helpful to know both representations. Both are often referenced and tabulated in many books and online.

The spherical harmonics  $Y_{j,m}(\theta,\phi)$  are mutually orthogonal. It is sometimes useful to normalize them as well. In V3p19-7, Feynman says there are several normalization conventions. He recommends:

$$Y_{j,m}(\theta,\phi) = \langle j,0 | R_y(\theta) R_z(\phi) | j,m \rangle \sqrt{\{(2j+1)/4\pi\}}$$

Normalizing  $Y_{j,m}(\theta,\phi)$  isn't important to us here, because we will ultimately normalize the entire wave function.

The  $Y_{j,m}$  functions for  $j=0, 1$ , and  $2$  are tabulated here.

| P | j | m | $Y_{j,m}(\theta,\phi)$                                                                                                         |
|---|---|---|--------------------------------------------------------------------------------------------------------------------------------|
| s | + | 0 | 0                                                                                                                              |
| p | - | 1 | $\begin{cases} +1 & -\sin\theta e^{+i\phi} / \sqrt{2} \\ 0 & \cos\theta \\ -1 & +\sin\theta e^{-i\phi} / \sqrt{2} \end{cases}$ |

| P | j | m | $Y_{j,m}(\theta,\phi)$                                                                                                                                                                                                                                            |
|---|---|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| d | + | 2 | $\begin{cases} +2 & +\sin^2\theta e^{+2i\phi} \sqrt{(3/8)} \\ +1 & +\sin\theta \cos\theta e^{+i\phi} \sqrt{(3/2)} \\ 0 & (3\cos^2\theta - 1)/2 \\ -1 & -\sin\theta \cos\theta e^{-i\phi} \sqrt{(3/2)} \\ -2 & -\sin^2\theta e^{-2i\phi} \sqrt{(3/8)} \end{cases}$ |

In these tables, the first column provides the symbol used for each value of  $j$ . These symbols originated well before anyone understood the underlying physics. While somewhat archaic, they remain in frequent use.

The second column, labeled P, is the parity of the electron's orbital state, which equals  $(-1)^i$ .

## General Solution for Hydrogen

Ignoring normalization, the form of the general solution that we have derived is:

$$\psi = Y_{j,m}(\theta, \phi) F_j(r) \exp\{-iEt/\hbar\}$$

We have solutions for  $Y_{j,m}(\theta, \phi)$ . Now let's find solutions for  $F_j(r)$ .

We return to the full polar coordinate version of the Schrödinger equation.

$$i\hbar d\psi/dt = -\hbar^2/2\mu \nabla^2\psi - (e^2/r) \psi$$

or

$$-(2\mu/\hbar^2) \{i\hbar d\psi/dt + e^2/r\} \psi = \nabla^2\psi$$

When we plug the general solution into this equation, the factor  $\exp\{-iEt/\hbar\}$  will be common to all terms; the only change it makes is replacing the time derivative term with  $E\psi$ . After cancelling the common exponential, we obtain:

$$-(2\mu/\hbar^2)(E+e^2/r)Y_{j,m}(\theta, \phi)F_j(r) = \nabla^2 Y_{j,m}(\theta, \phi)F_j(r)$$

$$\begin{aligned} -(2\mu/\hbar^2)(E+e^2/r) Y_{j,m} F_j &= Y_{j,m} \partial^2(rF_j)/\partial r^2 /r \\ &+ [F_j/(r^2 \sin\theta)] \partial(\sin\theta \partial Y_{j,m}/\partial\theta)/\partial\theta \\ &+ [F_j/(r^2 \sin^2\theta)] (\partial^2 Y_{j,m}/\partial\theta^2) \end{aligned}$$

We next move all the r-dependent terms to the left side of the equation by multiplying everything by  $r^2/F_j$  and rearranging.

$$\begin{aligned} -\{(r^2/F_j) [\partial^2(rF_j)/\partial r^2 /r + (2\mu/\hbar^2)(E+e^2/r) F_j]\} Y_{j,m} &= \\ [\partial(\sin\theta \partial Y_{j,m}/\partial\theta)/\partial\theta + (\partial^2 Y_{j,m}/\partial\theta^2)/\sin\theta] /(\sin\theta) \end{aligned}$$

Let's convert to natural units as we did in the spherically symmetric case.

$$\begin{aligned} \rho &= r / r_B = r/(\hbar^2/\mu e^2) \\ \varepsilon &= E / R_y = E/(\mu e^4/2\hbar^2) \end{aligned}$$

This yields:

$$\begin{aligned} -\{(\rho^2/F_j) [\partial^2(\rho F_j)/\partial \rho^2 / \rho + (\varepsilon+2/\rho) F_j]\} Y_{j,m} &= \\ [\partial(\sin\theta \partial Y_{j,m}/\partial\theta)/\partial\theta + (\partial^2 Y_{j,m}/\partial\theta^2)/\sin\theta] /(\sin\theta) \end{aligned}$$

In V3p19-10, Feynman makes two key points. Firstly, the right hand side (RHS) of this equation does not depend on the orbital radius  $\rho$  — its value is the same value for all values of  $\rho$  — which means

the left hand side (LHS) must also have the same value for all values of  $\rho$ . This seems remarkable because the LHS is full of  $\rho$ 's. Yet, after all the math is done, the value of the LHS cannot be a function of  $\rho$ . Please be sure to grasp this essential point — perhaps this will help:  $d(\text{RHS})/d\rho=0$  means  $d(\text{LHS})/d\rho=0$ .

The second key point is that the LHS term in  $\{ \}$ 's has no angular dependence except  $F_j$ , which depends on the total orbital angular momentum  $j$ . The combination of these two key points means the  $\{ \}$  term must be some quantity, call it  $K_j$ , that depends only on  $j$ , and not on  $m$ ,  $\rho$ ,  $\theta$ , or  $\phi$ .

We can therefore rewrite the above equation as:

$$-\{(\rho^2/F_j) [\partial^2(\rho F_j)\partial\rho^2/\rho + (\epsilon+2/\rho) F_j]\} Y_{j,m} \\ = -K_j Y_{j,m} = \\ [\partial(\sin\theta \partial Y_{j,m}/\partial\theta)/\partial\theta + (\partial^2 Y_{j,m}/\partial\phi^2)/\sin\theta] /(\sin\theta)$$

We now split this into two separate equations. First is the angular equation:

$$-K_j Y_{j,m} = \\ [\partial(\sin\theta \partial Y_{j,m}/\partial\theta)/\partial\theta + (\partial^2 Y_{j,m}/\partial\phi^2)/\sin\theta] /(\sin\theta)$$

Second is the radial equation:

$$-\{(\rho^2/F_j) [\partial^2(\rho F_j)\partial\rho^2/\rho + (\epsilon+2/\rho) F_j]\} Y_{j,m} \\ = -K_j Y_{j,m}$$

We rearrange the latter yielding:

$$\partial^2(\rho F_j)\partial\rho^2/\rho + (\epsilon+2/\rho) F_j = F_j K_j / \rho^2$$

Our plan is to use the angular equation to calculate  $K_j$ , and then use  $K_j$  and the radial equation to calculate  $F_j$ .

So, what is  $K_j$ ?

Since  $K_j$  does not depend on the z-component of orbital angular momentum  $m$ , we are free to evaluate the first equation for any value of  $m$ . We choose  $m=j$ . Let's calculate  $Y_{jj}$  (still ignoring normalization).

$$Y_{jj}(\theta,\phi) = \langle j,0 | R_y(\theta) R_z(\phi) | j,j \rangle$$

As we have often discussed,  $R_z(\phi)$  is particularly simple. In this case:

$$R_z(\phi) | j,j \rangle = \exp\{ij\phi\} | j,j \rangle$$

Next, the y-rotation. We recall (Chapter 17) that for a spin 1/2 entity:

$$\langle m=+1/2 | R_y(\theta) | m=+1/2 \rangle = +\cos(\theta/2)$$

$$\langle m=-1/2 | R_y(\theta) | m=+1/2 \rangle = -\sin(\theta/2)$$

A  $|j,j\rangle$  entity is equivalent to  $2j$  spin  $1/2$  entities that are all spin up. Projecting a rotated  $|j,j\rangle$  state onto  $\langle j,0|$  is equivalent to projecting  $j$  rotated spin  $+1/2$  states onto  $\langle +1/2|$  and projecting  $j$  rotated spin  $+1/2$  states onto  $\langle -1/2|$ . The amplitude for the spin  $1/2$  rotations is:

$$\begin{aligned} & [+ \cos(\theta/2)]^j [-\sin(\theta/2)]^j \\ &= [-\cos(\theta/2) \sin(\theta/2)]^j \\ &= [-(\sin\theta)/2]^j \end{aligned}$$

Putting the two rotation factors together, we get (ignoring normalization):

$$Y_{jj}(\theta,\phi) = [\sin\theta]^j \exp\{ij\phi\}$$

We now evaluate two pieces of the angular equation.

First  $\partial(\sin\theta \partial Y_{jj}/\partial\theta)/\partial\theta$  :

$$\sin\theta \partial Y_{jj}/\partial\theta = j[\sin\theta]^j \cos\theta \exp\{ij\phi\}$$

$$\begin{aligned} & \partial (\sin\theta \partial Y_{jj}/\partial\theta) / \partial\theta \\ &= \{j^2 [\sin\theta]^{j-1} \cos^2\theta - j[\sin\theta]^j \sin\theta\} \exp\{ij\phi\} \end{aligned}$$

Second  $\partial^2 Y_{jj}/\partial\phi^2$  :

$$\partial^2 Y_{jj}/\partial\phi^2 = [\sin\theta]^j (-j^2) \exp\{ij\phi\}$$

Now put the pieces together:

$$\begin{aligned} -K_j Y_{jj} &= \\ & \{\partial(\sin\theta \partial Y_{jj}/\partial\theta)/\partial\theta + (\partial^2 Y_{jj}/\partial\phi^2)/\sin\theta\} / (\sin\theta) \end{aligned}$$

$$\begin{aligned} -K_j Y_{jj} &= \{j^2 [\sin\theta]^{j-2} \cos^2\theta - j [\sin\theta]^j\} \exp\{ij\phi\} \\ &\quad - j^2 [\sin\theta]^{j-2} \exp\{ij\phi\} \end{aligned}$$

$$= j [\sin\theta]^{j-2} \{j \cos^2\theta - \sin^2\theta - j\} \exp\{ij\phi\}$$

$$= j [\sin\theta]^{j-2} \{-j \sin^2\theta - \sin^2\theta\} \exp\{ij\phi\}$$

$$= -j [\sin\theta]^j \{j + 1\} \exp\{ij\phi\}$$

$$= -j(j+1) Y_{jj}$$

$$K_j = j(j+1)$$

After all that,  $K_j$  is quite simple. We now plug  $K_j$  into the radial equation:

$$\partial^2(\rho F_j) \partial \rho^2 / \rho + (\epsilon + 2/\rho) F_j = F_j j(j+1) / \rho^2$$

$$\partial^2(\rho F_j) \partial \rho^2 = -\{\epsilon + 2/\rho - j(j+1)/\rho^2\} \rho F_j$$

This is the same equation that we had in the spherically symmetric case, except now we have the additional  $j(j+1)$  term on the right.

In V3p19-11, Feynman addresses the physical significance of this seemingly “mysterious” term. Consider the classical physics problem of an object of mass  $\mu$  orbiting a central potential, such as a planet orbiting a star. The conservation of total energy  $U$ , potential plus kinetic, can be written:

$$U = V(r) + \mu v^2/2 = \text{constant}$$

Now, we split the object’s velocity  $v$  into a radial component  $v_r$  and a tangential component  $r\omega$  ( $\omega = d\theta/dt$ ), yielding:

$$U = V(r) + \mu r^2 \omega^2/2 + \mu v_r^2/2$$

Next add the requirement that total angular momentum  $L$  is conserved.

$$L = \mu r^2 \omega = \text{constant}$$

$$L^2 / 2\mu r^2 = \mu^2 r^4 \omega^2 / 2\mu r^2 = \mu r^2 \omega^2/2$$

We plug this expression into the energy conservation equation.

$$U = V(r) + L^2/2\mu r^2 + \mu v_r^2/2$$

$$U = V^*(r) + \mu v_r^2/2$$

$$\text{with } V^*(r) = V(r) + L^2/2\mu r^2$$

Since  $L$  is constant,  $V^*$  is a function of  $r$  only. We can consider  $V^*$  to be the *effective potential* for orbits with angular momentum. For a given  $r$ , a certain amount of energy is tied up in the form of angular motion, and is therefore unavailable for radial motion. This is mathematically similar to the potential  $V(r)$ : for a given  $r$ , the  $V$  determines how much energy is available for motion.

In quantum mechanics, we found that the angular momentum term is:

$$j(j+1)/\rho^2$$

To convert this from natural units to conventional units, multiply by  $R_y r_b^2$  and get:

$$R_y r_b^2 j(j+1)/\rho^2 = (\mu e^4/2\hbar^2)(\hbar^4/\mu^2 e^4) j(j+1)/r^2$$

$$R_y r_b^2 j(j+1)/\rho^2 = j(j+1) \hbar^2/2\mu r^2$$

We see that quantum mechanics has the same angular momentum term, with  $L^2$  replaced by  $j(j+1)\hbar^2$ .

Let’s now solve the radial differential equation following the same procedure we used in the spherically symmetric case.

$$\partial^2(\rho F_j) \partial \rho^2 = -\{\varepsilon + 2/\rho - j(j+1)/\rho^2\} \rho F_j$$

Replace  $\rho F_j$  with  $g(\rho) \exp\{-\beta\rho\}$ . As before:

$$\begin{aligned} & \partial^2(g(\rho) \exp\{-\beta\rho\}) / \partial \rho^2 \\ &= [-2\beta \partial g / \partial \rho + \partial^2 g / \partial \rho^2 + \beta^2 g] \exp\{-\beta\rho\} \end{aligned}$$

$$\begin{aligned} & [-2\beta \partial g / \partial \rho + \partial^2 g / \partial \rho^2 + \beta^2 g] \\ &= -\{\varepsilon + 2/\rho - j(j+1)/\rho^2\} g \end{aligned}$$

$$\partial^2 g / \partial \rho^2 - 2\beta \partial g / \partial \rho + \{\beta^2 + \varepsilon + 2/\rho - j(j+1)/\rho^2\} g = 0$$

Again we set the arbitrary parameter  $\beta = \sqrt{-\varepsilon}$ , recalling that  $\varepsilon$  is negative for a bound electron. This reduces the equation somewhat.

$$\partial^2 g / \partial \rho^2 - 2\beta \partial g / \partial \rho + \{2/\rho - j(j+1)/\rho^2\} g = 0$$

Now let  $g(\rho) = \sum_k a_k \rho^k$ , with the summation being from  $k=0$  to  $k=\infty$ . The equation then becomes:

$$\begin{aligned} & \sum_k a_k \{k(k-1)\rho^{k-2} - 2\beta k\rho^{k-1} + 2\rho^{k-1} - j(j+1)\rho^{k-2}\} \\ &= 0 \end{aligned}$$

$$\begin{aligned} & \sum_k a_k \{k(k-1) - j(j+1)\} \rho^{k-2} \\ &= -\sum_k a_k \{-2\beta k + 2\} \rho^{k-1} \end{aligned}$$

As before, since the sums are independent, we may shift the summation indices independently in order to obtain the same power of  $\rho$  in each term. Replace the summation index  $k$ , with  $k+1$  on the left hand side.

$$\begin{aligned} & \sum_k a_{k+1} \{(k+1)k - j(j+1)\} \rho^{k-1} \\ &= \sum_k 2a_k \{\beta k - 1\} \rho^{k-1} \end{aligned}$$

The summation on the left side begins with  $k=-1$ , and on the right side begins with  $k=0$ . Examine the  $k=-1$  sum.

$$a_0 \{0 - j(j+1)\} \rho^{-2} = 0$$

For this to be true for all  $\rho$ ,  $a_0$  must equal 0. This means we can start both sums at  $k=0$ , and write:

$$0 = \sum_k [a_{k+1} \{k(k+1) - j(j+1)\} - 2a_k \{\beta k - 1\}] \rho^{k-1}$$

Now examine the  $k=0$  terms.

$$\begin{aligned} 0 &= [a_1 \{0 - j(j+1)\} - 2a_0 \{-1\}] \rho^{-1} \\ 0 &= a_1 j(j+1) \end{aligned}$$

Since we are solving for states with  $j>0$ , this means  $a_1=0$ .

For all other powers of  $\rho$ , the summation equation is valid for all  $\rho$  only if the expression in [ ]'s equals zero for all values of  $n$ . This means:

$$\begin{aligned} 0 &= a_{k+1} \{k(k+1) - j(j+1)\} - 2a_k \{\beta k - 1\} \\ a_{k+1} \{k(k+1) - j(j+1)\} &= 2a_k \{\beta k - 1\} \\ a_{k+1} &= 2a_k \{\beta k - 1\} / \{k(k+1) - j(j+1)\} \end{aligned}$$

This equation imposes two constraints.

The first constraint is the same as in the spherically symmetric case: to prevent wave functions that increase exponentially at large radius,  $\beta k$  must equal 1 for some value of  $k$ , which we will call  $n$ . Hence,  $\beta = 1/n$  for some integer  $n$ .

The second constraint is new. We found above that if  $j > 0$ ,  $a_1$  must be 0. This would make all other  $a$ 's equal zero, if not for the denominator. If the denominator were always non-zero, every  $a_{k+1}$  would be a finite multiple of  $a_k$  and would therefore be zero. This provides no solution at all for the electron wave function. However, the denominator does equal zero when  $k=j$ , making the equation meaningless for  $k=j$ . The equation thus imposes no requirement on  $a_{j+1}$ : it can be anything. And subsequent  $a$ 's can also be non-zero, as determined by the equation.

Therefore the sum over  $k$  is constrained to terms with  $k=j+1$  through  $k=n$ . Note that we have valid solutions only if  $j < n$ .

As before, since the only allowed values of  $\beta$  are  $1/n$  for some integer  $n$ , the energy of allowed electron orbits is once again given by:

$$\begin{aligned} 1/n &= \beta = \sqrt{-\epsilon} \\ 1/n &= \sqrt{-E_n/R_y} \\ E_n &= -R_y/n^2 \end{aligned}$$

All the pieces of the electron wave function are put together in the review section below.

## Proof of Polynomial Theorem

We prove here the theorem quoted above for a Taylor series. We represent a function  $f(x)$  by a polynomial summed from  $n=0$  to  $n=+\infty$ . If:

$$0 = f(x) = \sum_n a_n x^n, \text{ for all } x$$

the only solution is  $a_n = 0$  for all  $n$ .

For  $x=0$ , the equation reduces to  $0 = a_0$ . That's one down, infinitely more  $a$ 's to go.

Next take the derivative of  $f(x)$ , and evaluate that at  $x=0$ :

$$0 = df/dx = \sum_n n a_n x^{n-1}$$

$$0 = df(x=0)/dx = 1 \cdot a_1 \cdot 1 + \sum_{n>1} n a_n (0)^{n-1}$$

$$0 = a_1$$

The derivative of  $df/dx$ , evaluated at  $x=0$ , proves that  $a_2=0$ .

Successively evaluating higher order derivatives at  $x=0$  proves that  $a_n = 0$  for all  $n$ .

QED

## Chapter 29 Review: Key Ideas

• For an electron in a hydrogen atom in a spherically symmetric s-orbit, in a stationary state with time dependence  $\exp\{-iEt/\hbar\}$ , Schrödinger's equation in polar coordinates is:

$$\partial^2(r\psi)/\partial r^2 / r = - (2\mu/\hbar^2) (E + e^2/r) \psi$$

where  $r$  is the electron's radial coordinate, and  $e^2=q_e^2/4\pi\epsilon_0$ . We convert to natural units using:

$$\rho = r / r_B; r_B = \hbar^2/\mu e^2 = \text{Bohr radius}$$

$$\epsilon = E / R_y; R_y = \mu e^4/2\hbar^2 = 1 \text{ Rydberg}$$

We also make these substitutions:

$$\rho\psi = g(\rho) \exp\{-\beta\rho\}, \text{ with } \beta=\sqrt{-\epsilon}$$

$$g(\rho) = \sum_k a_k \rho^k, \text{ a Taylor series expansion}$$

Our differential equation becomes:

$$0 = \sum_k \{(k+1)ka_{k+1} - 2\beta ka_k + 2a_k\} \rho^{k-1}$$

To be valid for all  $\rho$ , the quantity in  $\{ \}$ 's must be zero for all  $n$ , resulting in discrete eigenstates with:

$$E_n = - R_y/n^2$$

$$\psi_n(\rho) = g_n(\rho) \exp\{-\rho/n\} / \rho$$

$$g_n(\rho) = \sum_k a_k \rho^k, \text{ sum over } k=1 \text{ to } k=n$$

$$a_{k+1} = -a_k \{2(1-k/n)/k(k+1)\}$$

• Orbits with angular momentum are not spherically symmetric, and require spherical harmonics defined by:

$$Y_{j,m}(\theta,\phi) = \langle j,0 | R_y(\theta)R_z(\phi) | j,m \rangle$$

We can separate the angular dependence  $Y_{j,m}(\theta,\phi)$  from the radial dependence  $F_j(r)$ , which is the

solution of:

$$\partial^2(\rho F_j) \partial \rho^2 = -\{\varepsilon + 2/\rho - j(j+1)/\rho^2\} \rho F_j$$

With the same substitutions used above, we find:

$$a_{k+1} = 2a_k \{\beta k - 1\} / \{k(k+1) - j(j+1)\}$$

This equation requires:  $\beta = 1/n$  for some integer  $n$ ;  $k > j$ ; and  $j < n$ .

• The electron wave function in a hydrogen atom, for quantum numbers  $n$ ,  $j$ , and  $m$ , is:

$$\Psi_{n,j,m}(\theta, \varphi, \rho, t) = \alpha Y_{j,m}(\theta, \varphi) F_{n,j}(\rho) \exp\{-iE_n t/\hbar\}$$

$$E_n = -R_y/n^2$$

$$\rho = r / r_b = r / (\hbar^2/\mu e^2)$$

$$a_{k+1} = 2a_k \{k/n - 1\} / \{k(k+1) - j(j+1)\}$$

$$F_{n,j}(\rho) = \exp\{-\rho/n\} \sum_k a_k \rho^{k-1}$$

Here  $\alpha$  is the normalization factor,  $\mu$  is the electron mass, and the sum over  $k$  goes from  $k=j+1$  to  $k=n$ .

# Chapter 30

## The Periodic Table of Elements

---

This chapter describes qualitatively the structure of the Periodic Table of Elements.

### Hydrogen Wave Functions Overview

Let's review the key features of the wave functions of electrons bound in hydrogen atoms that we derived in the prior chapter.

Nature severely restricts the allowed orbits of bound electrons, as described by quantum mechanics and Schrödinger's equation. We derived a discrete set of orbital basis states, each of which is characterized by three integers:

$n$  determines  $\langle 1/r \rangle$ , the orbit's average inverse radius. The *principal quantum number*  $n$  can have any integral value greater than zero. The maximum  $n$  found in known elements is 7.

$j$  specifies the orbit's total angular momentum:  $L=j\hbar$ . The *orbital angular momentum quantum number*  $j$  can have any integral value from 0 through  $n-1$ .

$m$  specifies the component of orbital angular momentum along any selected axis, typically the  $z$ -axis. The *magnetic quantum number*  $m$  can have any integral value from  $-j$  through  $+j$ .

The wave functions  $\psi_{n,j,m}(\theta,\phi,r,t)$  of these orbital basis states can be separated into these factors:

$$\psi = \alpha Y_{j,m}(\theta,\phi) F_{n,j}(r) \exp\{-iE_n t/\hbar\}$$

$\alpha$  provides normalization

$Y_{j,m}(\theta,\phi)$  is the angular dependence

$F_{n,j}(r)$  is the radial dependence

$\exp\{-iE_n t/\hbar\}$  is the time dependence

The energy of each basis state is given by:

$$E_n = -R_y/n^2$$

Here  $R_y = -13.6 \text{ eV}$  is the Rydberg constant.

Each orbital basis state, each unique set of the three quantum numbers  $n$ ,  $j$ , and  $m$ , can accommodate a maximum of two electrons, one with spin up and another with spin down.

## Hydrogen s-Orbits

The lowest energy state is the spherically symmetric 1s-orbit for which  $n=1$ ,  $j=0$ ,  $m=0$ . In V3p19-12, Feynman notes that this ground state is non-degenerate: there is only one orbital state with this energy.

Each higher value of  $n$  also has exactly one spherically symmetric s-orbit; these are named “ns”, as in 2s, 3s, etc. The radial dependence of each s-orbit is defined by  $F_{n,0}(r)$ , which equals an exponential multiplied by a polynomial in  $r$  of degree  $n-1$ . For example, the polynomial for the 3s-orbit is:

$$F_{3,0}(r) = \{1 - 2\rho/3 + 2\rho^2/27\} \exp(-\rho/3)$$

with  $\rho = r / r_B$

and Bohr radius  $r_B = 0.0529 \text{ nm}$

A polynomial of degree  $n-1$  has  $n-1$  roots, which means  $F_{n,0}(r)=0$  at  $n-1$  values of  $r$ . The 1s wave function decreases monotonically from  $r=0$ , and is never zero. The 2s wave function is zero at one value of  $r$ , the 3s wave function is zero at two values of  $r$ , etc. At every value of  $r$  for which  $F_{n,0}(r)=0$ , there is a spherical *nodal surface*;  $\psi=0$  at every point on the surface of the sphere with that radius.

The exponential radial factor declines less rapidly for larger values of  $n$ , which means the wave function extends to larger radii for larger  $n$ .

The first three s-orbits are illustrated in the left most column of Figure 30-1, where darker areas indicate higher probability densities.

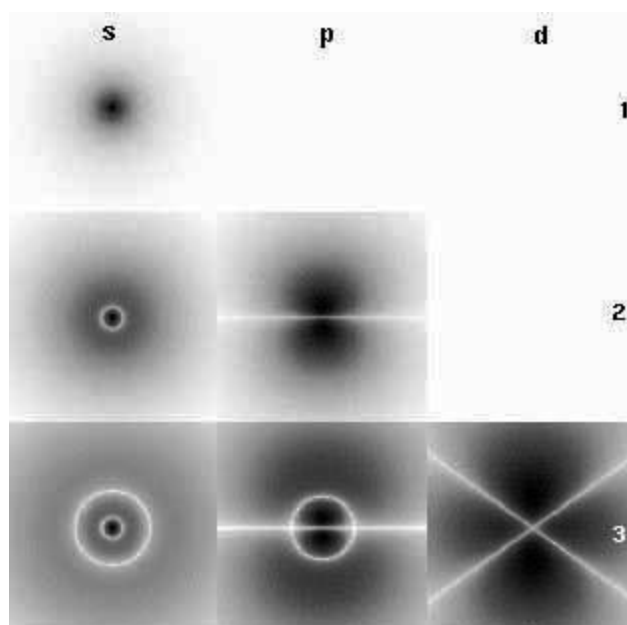


Figure 30-1 Hydrogen Orbits 1s Through 3d

## Hydrogen p- & d- Orbits

For each  $n > 1$ , there are exactly three p-orbits with  $j=1$  and  $m=+1, 0$ , and  $-1$ . And for each  $n > 2$ , there are exactly five d-orbits with  $j=2$  and  $m=+2, +1, 0, -1$ , and  $-2$ . The lowest-energy p- and d-orbits are illustrated in the middle and right columns of Figure 30-1, respectively.

The radial dependence of orbits with non-zero angular momentum also equals an exponential multiplied by a polynomial in  $r$ :

$$F_{n,j}(\rho) = \exp\{-\rho/n\} \sum_k a_k \rho^{k-1}$$

But the difference here is that the sum over  $k$  is from  $k=j+1$  to  $k=n$ , which excludes powers of the radius less than  $j$ . Hence, the polynomial in  $r$  still has degree  $n-1$ , and still has  $n-1$  roots, but  $j$  of those roots are at  $r=0$ .

This means all orbits with  $j > 0$  have  $\psi=0$  at  $r=0$ . This makes sense classically: an orbit that passes through the center has zero lever arm and therefore zero angular momentum. In quantum terms, if an orbit has a non-zero amplitude at  $r=0$ , its angular momentum  $j$  must be zero.

As  $j$  increases, more terms with low powers of  $r$  are excluded from the  $F_{n,j}(\rho)$  polynomial, increasingly skewing it toward higher powers of  $r$ . The result is that, as total angular momentum increases, the wave function is excluded from a progressively larger central zone surrounding the nucleus. This changes the wave function's shape, but does not change  $\langle 1/r \rangle$ , the average inverse radius, which is still determined only by  $n$ .

Eliminating low powers of  $r$  has another effect. The lowest energy p-orbit, 2p, has no spherical nodal surface. It is zero only at the center,  $\rho=0$ , because its polynomial has only one term:

$$F_{2,1}(\rho) = a_2 \rho \exp\{-\rho/n\}$$

Similarly, the lowest energy orbit with total angular momentum  $j$  has no spherical nodal surface for any value of  $j$ . Note in Figure 30-1 that state 3d also has no spherical nodal surface.

From the prior chapter, the angular dependence of each p-orbital basis state is:

$$|j,m\rangle = |1,+1\rangle: +\sin\theta \exp\{+i\phi\} / \sqrt{2}$$

$$|j,m\rangle = |1,0\rangle: +\cos\theta$$

$$|j,m\rangle = |1,-1\rangle: -\sin\theta \exp\{-i\phi\} / \sqrt{2}$$

The  $m=0$  p-orbital basis state has a *nodal plane*:  $\psi=0$  on the  $z=0$  plane, where  $\theta=\pi/2$ . The  $m=\pm 1$  p-orbital basis states are zero on the  $z$ -axis, where  $\theta=0$  or  $\theta=\pi$ .

In V3p19-13, Feynman emphasizes that what we have described are a complete set of basis states, and that any linear combinations of these states are also valid electron wave functions in hydrogen atoms.

For example, as shown above, the 2p  $m=0$  basis state is oriented along the z-axis, with a cosine dependence on the polar angle. Feynman asks about the existence of states “along the x-axis or along the y-axis. Are these perhaps the  $m=+1$  and  $m=-1$  states? No.”

Recall that the three  $m$ -states are stationary states of the same definite energy, with the same time dependence factor,  $\exp\{-iEt/\hbar\}$ , which is shared by any linear combination of these states. Feynman says the following  $|j,m\rangle$  combinations define p-orbits oriented along each axis:

z-axis:  $|1,0\rangle$

x-axis:  $\{|1,+1\rangle + |1,-1\rangle\}/\sqrt{2}$

y-axis:  $\{|1,+1\rangle - |1,-1\rangle\}/i\sqrt{2}$

The angular factors,  $Y_{j,m}(\theta,\phi)$ , of p-orbits involve sine and cosine functions to the first power. By contrast, d-orbits involve those functions to the second power. The angular dependence of each d-orbital basis state is:

$|j,m\rangle = |2,+2\rangle: +\sin^2\theta \exp\{+i2\phi\} \sqrt{3/8}$

$|j,m\rangle = |2,+1\rangle: +\sin\theta \cos\theta \exp\{+i\phi\} \sqrt{3/2}$

$|j,m\rangle = |2,0\rangle: +(3\cos^2\theta - 1) / 2$

$|j,m\rangle = |2,-1\rangle: -\sin\theta \cos\theta \exp\{-i\phi\} \sqrt{3/2}$

$|j,m\rangle = |2,-2\rangle: -\sin^2\theta \exp\{-i2\phi\} \sqrt{3/8}$

In general, orbits with total angular momentum  $j$  involve sinusoids to the  $j$ th power.

## Periodic Table of Elements

We will now apply our discoveries regarding single electron wave functions in hydrogen atoms to more complex elements with many electrons.

Each electron orbital state with quantum numbers  $|n,j,m\rangle$  may contain up to two electrons, one with spin up and the other spin down. Each group of  $|n,j\rangle$  states has  $2j+1$  different  $m$ -states that may contain up to  $4j+2$  electrons. The maximum number of electrons with principal quantum number  $n$ , the capacity of the  $n$ -shell, is:

Max in  $n = \sum_j (4j+2)$ , for  $j=0 \dots (n-1)$

Max in  $n = 4(\sum_j j) + (\sum_j 2)$ , for  $j=0 \dots (n-1)$

Max in  $n = 4 n(n-1)/2 + 2n$

Max in  $n = 2n^2 - 2n + 2n = 2n^2$

An atom with  $Z$  protons in its nucleus is defined to have *atomic number*  $Z$  and is called *element*  $\#Z$ . Generally, nuclei also contain some neutrons. All atoms of a given element must have the same number of protons, but they may have different numbers of neutrons, in which case they are said to be different *isotopes* of that element. Lead, for example, always has 82 protons, but lead has 38 known isotopes that contain from 96 to 133 neutrons.

If an atom with  $Z$  protons is electrically neutral, it must have  $Z$  electrons. Each electron is attracted by the nuclear charge  $+Z$ , but is also repelled by the negative charges of the  $Z-1$  other electrons.

To calculate the electron wave functions of an atom with  $Z$  protons and  $Z$  electrons, we could use a trial solution of the form:

$$\Psi = \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) \exp\{-iEt/\hbar\}$$

where  $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z$  are the position vectors of electrons #1 through # $Z$ . The Schrödinger equation for spherically symmetric states is:

$$\begin{aligned} E\psi = & \\ & -\hbar^2/2\mu \left\{ \sum_j \partial^2(\mathbf{r}_j\psi)/\partial r_j^2/r_j \right\} \\ & - \left\{ \sum_j Ze^2/r_j \right\} \psi \\ & + \left\{ \sum_{j>k} e^2/|\mathbf{r}_j - \mathbf{r}_k| \right\} \psi \end{aligned}$$

Here, the second line contains the sum of  $Z$  second order partial derivatives, one with respect to the radial coordinate of each electron. The third line contains the sum of the negative potential energy of each electron in the Coulomb field of the nucleus. The fourth line contains the sum of the positive potential energy of each pair of electrons due to their mutual Coulomb repulsion.

All this is the simplest case, without spin interactions and without angular momentum. Needless to say, no one has ever solved the complete multi-electron problem analytically.

It is possible, however, to gain some qualitative insights into how real electron wave functions vary with atomic number  $Z$ .

We will simplify the problem with several approximations. We assume each electron occupies the lowest available energy state. We will treat electrons as independent particles, subject to a few constraints. In accordance with the exclusion principle, at most two electrons can occupy each unique orbital state, each unique set of  $n$ ,  $j$ , and  $m$  quantum numbers. We begin with one electron in the Coulomb potential of a one-proton nucleus. We next add a second proton and a second electron. We assume the second electron experiences the potential of two protons *shielded* by a spherically symmetric charge density from the first electron. We continue in this way, adding protons to the nucleus and adding electrons to hydrogen-like  $|n,j,m\rangle$  states, until we reach  $Z$  electrons orbiting a nucleus of  $Z$  protons.

As we have seen, the radial distributions of electron wave functions grow more complex as  $n$  increases. As a result, shielding by “inner” electrons changes the effective Coulomb potential seen by “outer” electrons, making their effective potential more complex than simply  $1/r$ .

Let’s see how far these simplifying approximations can take us.

## **Z=1: Hydrogen**

This is the only case that is simple enough to solve analytically. We found that the ground state,  $1s$ ,

has energy  $E_1 = -13.60$  eV, which in chemical units is 1312 kJoules/mole. An atom's *ionization energy*,  $W_i$ , is the energy required to remove one electron. For hydrogen,  $W_i = -E_1$ . An element with a low ionization energy easily loses electrons in chemical reactions, while one with a high ionization energy does not.

## **Z=2: Helium**

Both electrons in helium occupy the 1s ground state (with opposite spins), and each partially shields the other from the nuclear charge. The ionization energy of helium is 24.59 eV. We might say this means each electron sees an effective potential of 1.8 protons ( $1.8 = 24.59/13.60$ ). Helium has the greatest ionization energy of any element. Also, since its electrons completely fill the  $n=1$  shell, a helium atom exerts no detectable attractive force on additional electrons. This combination means helium atoms never share electrons with other atoms and are therefore chemically inert.

## **Z=3: Lithium**

With the first two electrons filling the 1s ground state, the third electron must enter an  $n=2$  state. This raises an interesting question and some important physics: which angular momentum state within the  $n=2$  shell has the lowest energy?

In an atom with only one electron, that electron experiences the same  $-Ze^2/r$  potential everywhere. This is why the energies of hydrogen states depend only on their average inverse radius, which is solely determined by the principal quantum number  $n$ .

But in a multi-electron atom, electrons that are closer to the nucleus partially shield its attractive charge from electrons in more distant orbits. Due to this shielding, outer electrons experience a less negative potential and less binding energy. Electrons in s-states have a substantial amplitude to be at the atom's center, and thereby have some exposure to the full nuclear charge. But, electrons with orbital angular momentum have zero amplitude to be at the atom's center, as we discussed earlier. They therefore are more shielded by inner electrons and have less binding energy. The higher its orbital angular momentum becomes, the more an electron is shielded, and the less binding energy it has. Since binding energy is negative, electrons in high  $j$ -states have higher (less negative) energies.

This effect in multi-electron atoms is illustrated in Figure 30-2, where the horizontal axis is the principal quantum number  $n$ , and the vertical axis is indicative of which states have higher energy, but is not a true linear scale.

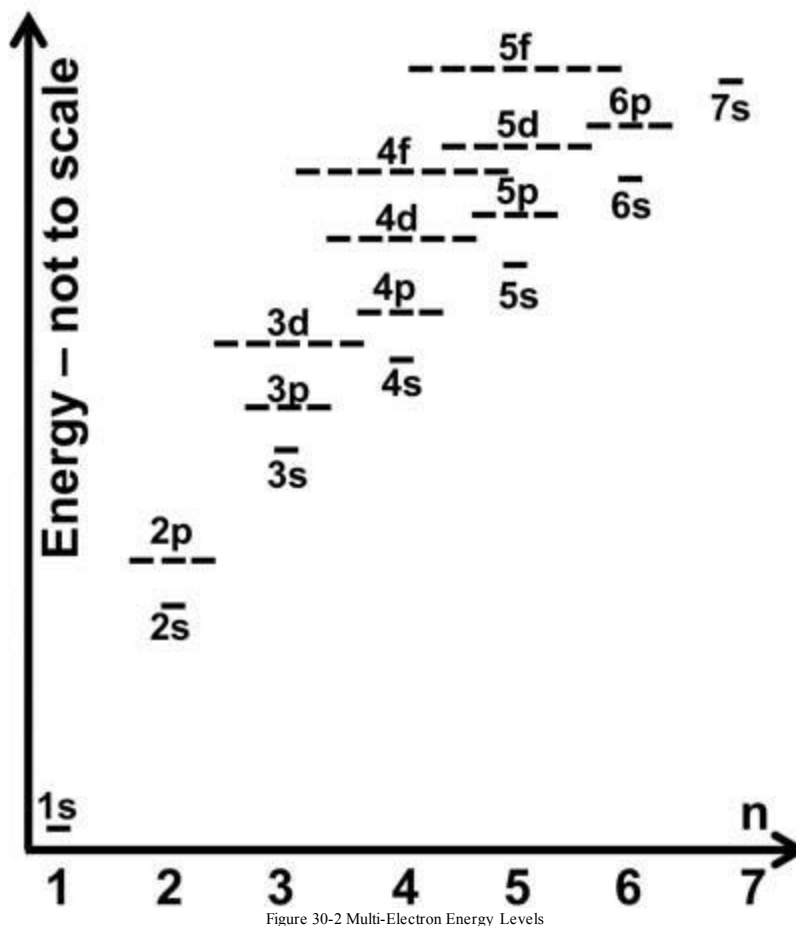


Figure 30-2 Multi-Electron Energy Levels

The key point is that energy levels in multi-electron atoms depend on both  $n$  and  $j$ , in sharp contrast to hydrogen-like energy levels.

The third electron of lithium, therefore, enters the 2s-orbit, not the higher energy 2p-orbit. The ionization energy of lithium is 5.39 eV, much less than hydrogen because 2s states have a larger average radius than 1s states.

### **Z=4: Beryllium**

The fourth electron completes the 2s state, giving beryllium a relatively high ionization energy: 9.32 eV.

### **Z=5 to 7: Boron to Nitrogen**

Electrons in the same  $|n,j,m\rangle$  state have the same orbital motion. This intensifies their mutual repulsion, reduces their binding energy, and increases their total energy. Therefore, electrons do not share the same  $|n,j,m\rangle$  state if a different state of the same orbital energy is available. This is why the next three electrons after beryllium enter three different 2p orbits. The ionization energies are: boron 8.30 eV; carbon 11.26 eV; and nitrogen 14.53 eV. Note that ionization energy steadily increases from  $Z=5$  to  $Z=7$ . The nuclear charge is increasing, while electron shielding changes slowly because the p-orbits are oriented in different directions.

### **Z=8 to 10: Oxygen to Neon**

Now that all 2p-orbits contain one electron, the next three electrons must fill out those orbits. This is

less energetically favorable, as evidenced by oxygen's ionization energy of 13.62 eV, which is 0.91 eV less than nitrogen despite oxygen having a greater nuclear charge. The ionization energy of fluorine is 17.42 eV, and of neon is 21.56 eV. Neon completes the  $n=2$  shell, making it a *noble gas* and chemically inert, just like helium.

### **Z=11 to 18: Sodium to Argon**

The 11th and 12th electrons fill the 3s state. The ionization energy for sodium is only 5.14 eV, and for magnesium is 7.65 eV. The 13th, 14th, and 15th electrons enter three different 3p-orbits. The ionization energies are: aluminum 5.99 eV; silicon 8.15 eV; and phosphorus 10.49 eV. The 16th, 17th, and 18th electrons fill the 3p-orbits. The ionization energies are: sulfur 10.36 eV; chlorine 12.97 eV; and argon 15.76 eV.

Comparing the sequence of elements lithium-to-neon with sodium-to-argon, the pattern of electron orbits and ionization energies is quite similar. This explains the similarity of chemical properties of corresponding elements: lithium to sodium; carbon to silicon; fluorine to chlorine; and neon to argon. This is why these pairs of atoms are placed in the same columns of the Periodic Table.

### **Z=19 and 20: Potassium and Calcium**

After 3p, we might well expect the next ten electrons to enter the five 3d-orbits (3d does not mean 3D; all orbits are three dimensional). But the 3d-orbits offer some surprises. Looking again at the energy level diagram, enlarged below in Figure 30-3.

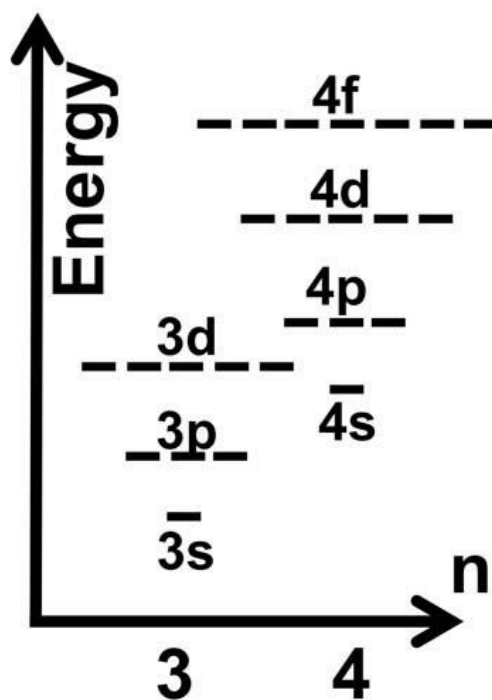


Figure 30-3 Enlargement of 3d, 4s Levels

Figure 30-3 shows that the energy of 3d-orbits is very slightly higher than the energy of 4s-orbits, confirming the increase in energy of higher angular momentum orbits in multi-electron atoms.

The 19th and 20th electrons, therefore, enter 4s rather than 3d-orbits, even though 4s has a higher value of  $\langle 1/r \rangle$  and is in that sense “outside” the 3d-orbit. Both types of orbits are distributed over a

wide range of  $r$ , as illustrated by Figure 30-4, hence “inside” and “outside” describe only average radii. For  $Z=19$  and  $Z=20$ , the ionization energies are: potassium 4.34 eV; and calcium 6.11 eV.

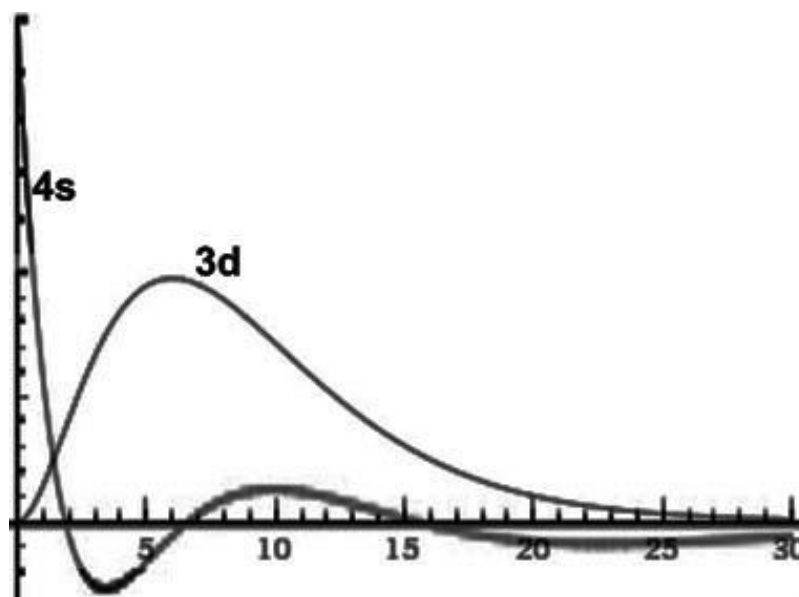


Figure 30-4 4s & 3d Radial Distributions

This figure compares the unnormalized  $F_{n_j}(\rho)$  for 4s and 3d hydrogen-like orbits, from  $\rho=0$  to  $\rho=30$  Bohr radii. The average inverse radii,  $\langle 1/r \rangle$ , for 3d and 4s-orbits correspond to 9 and 16 Bohr radii, respectively. However, 4s-orbits have substantial amplitudes near  $\rho=0$ , whereas 3d-orbits do not. Electrons in  $n=1$  and  $n=2$  shells more effectively shield 3d-orbits than 4s-orbits.

### **Z=21 to 30: Scandium to Zinc**

The next three electrons enter 3d-orbits. The ionization energies are: scandium 6.56 eV; titanium 6.83 eV; and vanadium 6.75 eV. As the 21st, 22nd, and 23rd electrons enter 3d-orbits, they increasingly shield the “outer” 4s electrons from the nuclear charge.

When the 24th electron arrives, it and the 20th electron both enter the last vacant 3d-orbits, leaving only one electron in the 4s-orbit. Filling a shell is very advantageous; even half-filling a subshell is advantageous.

The 25th electron refills the 4s state; this is manganese. The next three electrons (corresponding to iron, cobalt, and nickel) complete three previously half-occupied 3d-orbits.

The 29th electron (copper) repeats the pattern set by the 24th. It and the 25th electron both complete the 3d shell, leaving one electron in the 4s-orbit. The 30th electron refills the 4s state; this is zinc.

The ionization energies are: chromium 6.77 eV; manganese 7.43 eV; iron 7.90 eV; cobalt 7.88 eV; nickel 7.64 eV; copper 7.73 eV; and zinc 9.39 eV.

Since the 3d-orbits are closer to the nucleus than the 4s-orbits, as electrons fill 3d, the state of the outermost electrons changes very little. This explains why the chemical properties of these elements are so similar.

We have now discussed the aspects of electron orbits of most interest to physicists, so we won't

continue through the remaining 70+ elements.

## Topology of Molecules

Our analysis of the quantum states of electrons in atoms also explains the topology of molecules.

Oxygen, for example, has four 2p electrons. This means one p-orbit is filled and two are half-filled. The half-filled orbits allow oxygen to form chemical bonds with two other atoms. The p-orbits of an isolated oxygen atom are mutually orthogonal (a  $90^\circ$  angle between each pair), but that can change when atoms form molecules. In water, two hydrogen atoms bond with one oxygen atom. The two hydrogen electrons are more strongly attracted to the oxygen nucleus than to the hydrogen nuclei. While not completely removed from the hydrogen atoms, these electrons enter oxygen's half-filled p states, giving the oxygen atom a net negative charge and the hydrogen atoms a net positive charge. The hydrogen atoms push each other apart, due to their net positive charges. This effect stretches the bond angle to  $104.5^\circ$  from the intrinsic  $90^\circ$  angle between p-orbits.

The same effect occurs in  $\text{H}_2\text{S}$  and  $\text{H}_2\text{Se}$ . Sulfur and selenium have electron orbital structures similar to oxygen; sulfur has four 3p electrons, and selenium has four 4p electrons. Since both sulfur and selenium are larger atoms than oxygen, the two hydrogen atoms with which they bond are farther apart, reducing their repulsive force. The hydrogen bond angle is  $92.1^\circ$  in hydrogen sulfide, and  $91^\circ$  in hydrogen selenide.

Similar logic applies to nitrogen, which has three 2p electrons that half-fill each of the three 2p-orbits. In an isolated nitrogen atom, these orbits would be mutually orthogonal. In ammonia,  $\text{NH}_3$ , which we studied in Chapters 10 and 11, the hydrogen atoms again have a net positive charge and thereby repel one another. This stretches the hydrogen bond angle to  $107.8^\circ$ . Once again, larger atoms with the same electron structures have the same trend. Phosphorus has three 3p electrons and the hydrogen bond angle in  $\text{PH}_3$  is  $93.5^\circ$ . Arsenic has three 4p electrons and the hydrogen bond angle in  $\text{AsH}_3$  is  $91.8^\circ$ . Warning: don't measure these bond angles at home. While both phosphine and arsine are commonly used in semiconductor fabrication, both are extremely toxic.

## Chapter 30 Review: Key Ideas

- Each state with a unique set of quantum numbers  $n$ ,  $j$ , and  $m$ , may contain up to two electrons, one spin up and the other spin down. Each group of  $|n,j\rangle$  states has  $2j+1$  different  $m$ -states that may contain up to  $4j+2$  electrons. The maximum number of electrons with principal quantum number  $n$ , the capacity of the *n-shell*, equals  $2n^2$ .
- All orbits with  $j>0$  have zero amplitude at  $r=0$ . As  $j$  increases, the  $F_{n,j}(\rho)$  polynomial is increasingly skewed toward higher powers of  $r$ , excluding the wave function from a progressively larger central zone surrounding the nucleus. This changes the wave function's shape. but does not change  $\langle 1/r \rangle$ , its average inverse radius, which is still determined only by  $n$ .

- Orbital stationary states with the same  $n$  and  $j$  have the same energy. Any linear combination of such states is also a stationary state with the same energy. Feynman says the following combinations define p-orbits oriented along each axis:

z-axis:  $|1,0\rangle$

x-axis:  $\{|1,+1\rangle + |1,-1\rangle\}/\sqrt{2}$

y-axis:  $\{|1,+1\rangle - |1,-1\rangle\}/i\sqrt{2}$

- Higher angular momentum orbits in multi-electron atoms are shifted to higher energies, because they are more effectively shielded from the nuclear charge by “inner” electrons.
- In general, an orbit with total angular momentum  $j$  has an angular dependence that involves sinusoids raised to the  $j$ th power.

In V19-18, Feynman says: “The Schrödinger equation has been one of the great triumphs of physics. By providing the key to the underlying machinery of atomic structure it has given an explanation for atomic spectra, for chemistry, and for the nature of matter.” It also explains the structure of the Periodic Table of Elements.

# Chapter 31

## Quantum & Algebraic Operators

---

In V3p20-1, Feynman says:

“There are many ways of approaching the subject of quantum mechanics, and most books use a different approach from the one we have taken. As you go on to read other books you might not see right away the connections of what you will find in them to what we have been doing. Although we will also be able to get a few useful results, the main purpose of this chapter is to tell you about some of the different ways of writing the same physics. Knowing them you should be able to understand better what other people are saying.”

In the early days of classical physics, people wrote out all the x-, y-, and z-components of every natural law. Some people still do. Most however, now write these equations using vectors, making the equations simpler and more compact. While it is true that to obtain practical answers one must eventually enter component values, vector notation unquestionably simplifies complex mathematics.

Similarly, we have used operators and state vectors to make quantum mechanical equations less cluttered and make their meaning more evident.

A state vector  $|\psi\rangle$  has mathematical similarities to geometric vectors, but is more abstract, being a label for the physical condition of a quantum entity. Using state vectors, the natural laws of the quantum realm can be written as algebraic equations. Feynman provides the example of representing any state  $|\psi\rangle$  as a linear combination of basis states:

$$|\psi\rangle = \sum_j C_j |j\rangle$$

Here,  $C_j = \langle j|\psi\rangle$  is the amplitude for state  $|\psi\rangle$  to be in state  $|j\rangle$ , and the set of states  $|j\rangle$ , for  $j=1\dots N$ , forms a complete orthonormal basis.

We have used operators to describe how states change when they are rotated, projected onto a coordinate system with inverted axes, or simply left undisturbed to evolve over time. For example:

$$|\phi\rangle = A |\psi\rangle$$

represents operator  $A$  changing state  $|\psi\rangle$  into state  $|\phi\rangle$ . Feynman says the *definition* of operator  $A$  is given by the algebraic equation:

$$\langle k|\phi\rangle = \sum_j \langle k|A|j\rangle \langle j|\psi\rangle$$

Here  $\langle j|\psi\rangle$  is the “amount” of state  $|\psi\rangle$  in basis state  $|j\rangle$  — the amplitude of  $|\psi\rangle$  being in  $|j\rangle$ , or equivalently, the projection of  $|\psi\rangle$  onto  $|j\rangle$ . Similarly,  $\langle k|\phi\rangle$  is the “amount” of basis state  $|k\rangle$  in state  $|\phi\rangle$ . The complex number  $\langle k|A|j\rangle$  defines how much basis state  $|j\rangle$  is changed into basis state  $|k\rangle$  by operator  $A$  — the projection of  $A|j\rangle$  onto  $|k\rangle$ . The set of amplitudes  $\langle k|A|j\rangle$  are the components of a matrix that represents the operator  $A$ , according to:

$$A_{kj} = \langle k|A|j\rangle$$

So, Feynman says, the *operator* equation  $|\phi\rangle = A|\psi\rangle$  is a more abstract, more “high-class”, way of writing the *algebraic* equation  $\langle k|\phi\rangle = \sum_j \langle k|A|j\rangle \langle j|\psi\rangle$ .

The algebraic equation is expressed in one particular basis set, any set we wish to select. The operator equation is universal, independent of basis states. The distinction is perhaps more philosophical than practical, since obtaining practical results requires defining basis states. The operator-algebraic distinction is similar to the distinction between:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b}$$

and

$$\begin{aligned} c_x &= a_y b_z - a_z b_y \\ c_y &= a_z b_x - a_x b_z \\ c_z &= a_x b_y - a_y b_x \end{aligned}$$

Once one knows the components of a vector, matrix, or operator in any basis, one can calculate its components in any other basis using a suitable transformation.

In V3p20-2, Feynman says that the operator methodology opens new opportunities. An operator  $A$  might produce a new state  $A|\psi\rangle$  that may be unrelated to physical reality, but nonetheless useful analytically.

Recall some of the operators we have used in this course: the rotation operator  $R_z(\theta)$ ; the parity operator  $P$ ; and the Pauli spin operators  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$ .

We also defined in Chapter 23 the angular momentum operator  $J_z$  in terms of a rotation about the  $z$ -axis by an infinitesimal angle  $\epsilon$ :

$$R_z(\epsilon) = 1 + (i/\hbar) \epsilon J_z$$

or

$$R_z(\epsilon) |\psi\rangle = |\psi\rangle + (i/\hbar) \epsilon J_z |\psi\rangle$$

Both of the above equations express the same physics.

Rearranging the prior equation yields:

$$J_z |\psi\rangle = (\hbar/i\varepsilon) \{R_z(\varepsilon) |\psi\rangle - |\psi\rangle\}$$

Thus  $J_z |\psi\rangle$  is a state that is proportional to the difference between two other states:  $R_z(\varepsilon)|\psi\rangle$  and  $|\psi\rangle$ .

Also in Chapter 23, we defined the x-momentum operator  $p_x$  in terms of the displacement operator  $D_x(\delta)$ , which displaces the x-axis by  $+\delta$ , thereby subtracting  $\delta$  from all x-coordinate values.

$$D_x(\delta) = 1 + (i/\hbar) \delta p_x$$

$$D_x(\delta) |\psi\rangle = |\psi\rangle + (i/\hbar) \delta p_x |\psi\rangle$$

Quantum operators act on quantum states. By contrast, algebraic operators act on functions. For example, the operator  $d/dx$  can act on function  $f(x)$  to produce  $df(x)/dx$ , the x derivative of  $f$ .

We use both types of operators in quantum mechanics, often in the same equations. Sometimes the same notation is used for both. In truth, the distinction is somewhat pedantic.

In this chapter only, we will for clarity underline quantum operators and use bold italics for algebraic operators that also have quantum operator counterparts. Standard mathematical operators such as  $d/dx$  will not be changed. For example:

Quantum vs. Algebraic

$$\underline{H} |\psi\rangle \text{ vs. } \boldsymbol{H} \psi(x)$$

$$\underline{p}_x |\psi\rangle \text{ vs. } \boldsymbol{P}_x \psi(x)$$

One final general comment on operators: some quantum operators have the special property of being *Hermitian*. Consider the complex conjugates of the components of a quantum operator:

$$A_{ij} = \langle k | \underline{A} | j \rangle$$

$$(A_{ij})^* = (\langle k | \underline{A} | j \rangle)^*$$

$$(A_{ij})^* = \langle k | \underline{Z} | j \rangle, \text{ for some } \underline{Z}$$

$$(A_{ij})^* = \langle j | \underline{Z} | k \rangle$$

$$(A_{ij})^* = Z_{jk}$$

This says the  $jk$ -component of operator  $\underline{Z}$  equals the complex conjugate of the  $kj$ -component of operator  $\underline{A}$ . Operator  $\underline{Z}$  is called the *Hermitian adjoint* of operator  $\underline{A}$ . In quantum mechanics, many important operators have the special property that  $\underline{Z}=\underline{A}$ . Such operators are called *Hermitian*, or *self-adjoint*.

## Average Energy in Discrete Basis

Since measurement is a cornerstone of science, we now address the process of measuring the properties of a quantum system. Specifically, we consider measuring the energy of an atom that has a discrete set of basis states.

We have extensively discussed stationary states, states with definite energy. Imagine a situation in which a large number of atoms are known to exist in the same stationary state. An example might be hydrogen atoms at such a low temperature that  $kT$  is much less than the energy of the lowest excited state. Such atoms would be extremely unlikely to have any energy other than  $E_0$ , the ground state energy of hydrogen. If we measure the energy of one hydrogen atom after another, we will presumably find that each atom has energy  $E_0$ , and therefore, their average energy is also  $E_0$ .

Now consider a more interesting situation of a system in state  $|\psi\rangle$  that is not a stationary state of definite energy. An example would be hydrogen atoms at a very high temperature. We know that any state can be represented by a linear combination of eigenstates, basis states of definite energy. Let  $E_j$  be the energy of eigenstate  $|j\rangle$ , for  $j=0, 1, \dots$ . We then have:

$$|\psi\rangle = \sum_j C_j |j\rangle$$

If we measure the energy of each of a large number of atoms in state  $|\psi\rangle$ , we will sometimes measure  $E_0$ , sometimes  $E_1$ , etc. If an atom is measured to have energy  $E_j$ , this means that atom was in state  $|j\rangle$  when it was measured. In fact, the probability of measuring  $E_j$  is identically equal to the probability of the atom being in state  $|j\rangle$ , which is:

$$\text{Prob}(E_j) = \text{Prob}(j) = |C_j|^2$$

Note that if the states  $|j\rangle$  form a complete basis set, as every proper basis set should, we will never measure an energy that is not one of the eigenvalues  $E_j$ .

I pause here to mention a very important issue that Feynman chooses not to address in this course. This issue is called the “**measurement problem**” or “**wave function collapse**.” Quantum mechanics says that a state like  $|\psi\rangle$  is a superposition of basis states  $|j\rangle$ , with amplitudes to be in each of many different states simultaneously. If undisturbed,  $|\psi\rangle$  is not in one basis state **or** another, but is in **all** these states simultaneously. However, when state  $|\psi\rangle$  suffers a substantial external interaction, such as a measurement,  $|\psi\rangle$  instantaneously “collapses” to one and only one eigenstate of that measurement. In this case, the eigenstates of definite energy are the states  $|j\rangle$ . This means that, when we measure the energy of  $|\psi\rangle$ , one  $C_j$  becomes 1 and all other  $C_j$ ’s become zero, and all this occurs instantaneously throughout all of space. No one has yet provided a detailed physical explanation for exactly how this “magic” occurs. We will discuss this further in Chapter 34 of this eBook. Meanwhile, back to Feynman’s lecture.

After making  $M$  energy measurements, we obtain the average energy of state  $|\psi\rangle$  by simply summing the measured energy values and dividing by  $M$ . Alternatively, for each value of  $j$ , we can count how many times we measure the value  $E_j$ , call that number  $M_j$ , and then obtain the average energy as follows:

$$\begin{aligned} \langle E_\psi \rangle &= \{ \sum_j M_j E_j \} / \sum_j M_j \\ \langle E_\psi \rangle &= \{ \sum_j M_j E_j \} / M \\ \langle E_\psi \rangle &= \sum_j (M_j/M) E_j \end{aligned}$$

For very large M,  $M_j/M$  approaches  $\text{Prob}(j)$ ; this is what mathematicians mean by “probability.” In any real experiment, M is a finite number and there will be statistical fluctuations: the  $M_j/M$ ’s will differ somewhat from the  $\text{Prob}(j)$ ’s. Yet, in all normal circumstances, the mean value — the *expectation value* — of  $M_j/M = \text{Prob}(j)$ .

A general principle of statistics pertains to measurements of any physical quantity. If the measurements are independent samplings of a uniform population (such as atoms in the same state),  $\sigma_{mj}$ , the root-mean-square (rms) deviation of  $M_j$ , equals:

$$\sigma_{mj} = \sqrt{\{ M p(1-p) \}}, \text{ with } p = \text{Prob}(j)$$

When there are many probable outcomes and the probability of any particular outcome is small ( $p \ll 1$ ), the above simplifies to

$$\sigma_{mj} = \sqrt{\{ M_j \}}$$

This is a very important result, well worth remembering. When an experiment detects 100 Higgs bosons, or 100 photons from a galaxy 13 billion light-years away, the rms statistical uncertainty is  $\sqrt{100} = 10$ .

In V3p20-4, Feynman says with delight: “Now watch this trickery.”

For our quantum state  $|\psi\rangle$ :

$$\langle E_\psi \rangle = \sum_j (M_j/M) E_j = \sum_j |C_j|^2 E_j$$

$$\langle E_\psi \rangle = \sum_j E_j C_j^* C_j$$

$$\langle E_\psi \rangle = \sum_j E_j \langle \psi | j \rangle \langle j | \psi \rangle$$

$$\langle E_\psi \rangle = \langle \psi | \{ \sum_j E_j | j \rangle \langle j | \psi \rangle \}$$

or

$$\langle E_\psi \rangle = \langle \psi | \hat{H} \rangle$$

$$\text{with } |\hat{H}\rangle = \sum_j E_j | j \rangle \langle j | \psi \rangle$$

Since  $|j\rangle$  is an eigenstate, we know:

$$\underline{H} | j \rangle = E_j | j \rangle$$

Inserting this into the prior equation yields:

$$\langle E_\psi \rangle = \langle \psi | \{ \sum_j \underline{H} | j \rangle \langle j | \psi \rangle \}$$

$$\langle E_\psi \rangle = \langle \psi | \{ \underline{H} | \psi \rangle \}$$

$$\langle E_\psi \rangle = \langle \psi | \underline{H} | \psi \rangle$$

Remarkably, the eigenstates of definite energy have been eliminated from our equation. We don’t need to know a system’s eigenstates, or their energy levels, in order to measure that system’s average

energy. We only need to know the Hamiltonian components  $H_{jk}$  in terms of any basis states. We can then write:

$$\begin{aligned}\langle E_\psi \rangle &= \sum_{jk} \langle \psi | j \rangle \langle j | \underline{H} | k \rangle \langle k | \psi \rangle \\ \langle E_\psi \rangle &= \sum_{jk} \langle \psi | j \rangle H_{jk} \langle k | \psi \rangle\end{aligned}$$

Feynman says we can check the last equation for the special case of basis states that are stationary states, where the Hamiltonian is diagonal and  $\underline{H}|k\rangle = E_k|k\rangle$ .

$$\begin{aligned}\langle E_\psi \rangle &= \sum_{jk} \langle \psi | j \rangle \delta_{jk} E_k \langle k | \psi \rangle \\ \langle E_\psi \rangle &= \sum_j \langle \psi | j \rangle E_j \langle j | \psi \rangle\end{aligned}$$

This confirms the earlier equation.

This is a general result that is true for any quantum operator. We can, for example, replace the energy operator  $\underline{H}$  with the z-component of angular momentum operator  $\underline{L}_z$ .

$$\langle L_{z\psi} \rangle = \langle \psi | \underline{L}_z | \psi \rangle$$

Feynman says one could prove this by identifying a situation in which a system's energy is proportional to the z-component of its angular momentum. Since the prior equation holds for any form of energy in any situation, it also holds for the z-component of its angular momentum.

The general statement is: for any physically observable quantity  $A$  associated with a quantum operator  $\underline{A}$ , the average value of  $A$  for any state  $|\psi\rangle$  is:

$$\langle A_\psi \rangle = \langle \psi | \underline{A} | \psi \rangle$$

## Average Energy in Continuous Basis

Next we consider measuring the energy of an atom represented in a continuous set of basis states. Specifically, we will consider a one-dimensional coordinate basis  $|x\rangle$  and a one-dimensional state  $|\psi\rangle$ , where by definition  $\langle x | \psi \rangle = \psi(x)$ .

The prior summation equation employing discrete basis states was:

$$\langle E_\psi \rangle = \sum_{jk} \langle \psi | j \rangle \langle j | \underline{H} | k \rangle \langle k | \psi \rangle$$

We now replace the double summation with a *double integral*, integrating over two continuous bases:

$$\langle E_\psi \rangle = \iint \langle \psi | x \rangle \langle x | \underline{H} | X \rangle \langle X | \psi \rangle dx dX$$

Whereas a single integral is a sum along a line, a double integral is a sum across a 2D-plane — the  $x$ - $X$  plane in this case. One evaluates double integrals in two steps: (1) integrate over one variable; then (2) integrate over the other variable, both in the normal manner. It makes no difference which variable

is integrated first.

Here we assume  $|\psi\rangle$  and  $\psi(x)$  are normalized to yield a total probability of 1.

We can rewrite the prior equation as:

$$\langle E_\psi \rangle = \int \langle \psi | x \rangle \langle x | \hat{H} | \psi \rangle dx$$

with

$$\langle x | \hat{H} | \psi \rangle = \int \langle x | \hat{H} | X \rangle \langle X | \psi \rangle dX$$

Note that:

$$\langle \psi | x \rangle = \langle x | \psi \rangle^* = [\psi(x)]^* = \psi^*(x)$$

Also note that  $\langle x | \hat{H} | X \rangle$  is the continuous equivalent of the discrete  $\langle j | \hat{H} | k \rangle$ .

From Chapter 22, we use Dirac's delta function to collapse the integral over  $X$ . We also use Schrödinger's equation from that chapter to obtain:

$$\langle x | \hat{H} | \psi \rangle = \hat{H}\psi(x) = \left\{ -\hbar^2/2m d^2/dx^2 + V(x) \right\} \psi(x)$$

Define the Hamiltonian algebraic operator  $\hat{H}$  as:

$$\hat{H} = -\hbar^2/2m d^2/dx^2 + V(x)$$
$$\langle E_\psi \rangle = \int \psi^*(x) \hat{H} \psi(x) dx$$

Feynman notes that in the upper line, potential energy  $V$  is effectively a number, and all numbers can be considered multiplicative operators:  $V(x)\{\psi(x)\} = V(x) \times \psi(x)$ .

The Hamiltonian operators  $\hat{H}$  and  $\hat{H}$  are both energy operators, but with a difference. The former is an algebraic operator that acts on a function to produce a new function:  $\hat{H}\psi(x) = \hat{H}\psi(x)$ . The latter is a quantum operator that acts on a state to produce a new state:  $\hat{H}|\psi\rangle = |\hat{H}\psi\rangle$ . The former assumes a coordinate representation that, which can be any representation we choose. The latter makes no reference to any particular representation, but to solve any problem, some representation must be eventually chosen. In my view, the difference is modest.

Feynman notes that if  $\psi(x)$  is not normalized, the prior equation should be written:

$$\langle E_\psi \rangle = \int \psi^*(x) \hat{H} \psi(x) dx / \int \psi^*(x) \psi(x) dx$$

Feynman comments on the similarity between two forms of the same relationship.

$$\langle E_\psi \rangle = \langle \psi | \hat{H} | \psi \rangle$$

$$\langle E_\psi \rangle = \int \psi^*(x) \hat{H} \psi(x) dx$$

He says these two forms arise in coordinate representations involving *local* operators. He says an operator is local if  $\hat{A}$ , its quantum operator form, and  $A$ , its algebraic operator form, are related by:

$$\hat{A} \psi(x) = \int \langle x | \hat{A} | X \rangle \langle X | \psi \rangle dX$$

Some operators are local and some are not.

The one-dimensional integrals are easily extended to three dimensions as follows:

$$\langle E_\psi \rangle = \iiint \psi^*(\mathbf{r}) \hat{H} \psi(\mathbf{r}) dx dy dz$$

$$\hat{H} = -\hbar^2/2m \{d^2/dx^2 + d^2/dy^2 + d^2/dz^2\} + V(\mathbf{r})$$

$$\text{provided } \iiint \psi^*(\mathbf{r}) \psi(\mathbf{r}) dx dy dz = 1$$

To extend these equations to multi-particle systems: replace vector  $\mathbf{r}$  by a set of vectors, one for the position of each particle; and replace the three second-order partial derivatives by the sum of three such derivatives for each particle's coordinates.

The first of the above three equations enables us to calculate the energy of any system, provided we know its wave function. But, remarkably this equation can be useful even if we don't know the wave function.

In V3p20-7, Feynman gives an interesting real-life example. No one can solve the Schrödinger equation for the ground state of helium, because there are simply too many complex terms. We can however make some educated guesses as to what the wave function might be.

We can use a trial wave function  $|\psi\rangle$  and the above equation to compute  $\langle E_\psi \rangle$ . Unless we miraculously guess the true wave function, our trial  $\langle E_\psi \rangle$  will be higher than the true ground state energy. (To understand why, think of  $|\psi\rangle$  as a linear combination of states  $|j\rangle$  of definite energy, including the true ground state  $|0\rangle$ . If  $|\psi\rangle$  is not exactly equal to  $|0\rangle$ , our trial  $\langle E_\psi \rangle$  will contain contributions from higher energy states that are greater than  $E_0$ .)

Next, we can try a second possible solution  $|\phi\rangle$ . Whichever solution gives a lower energy is probably closer to the real solution. We can repeat this process, hoping to get ever closer to the true ground state  $|0\rangle$ .

Clever theoretical physicists (many are very clever indeed) use trial solutions with variable parameters, say  $p$  and  $q$ , thereby obtaining  $\langle E_\psi \rangle$  as a function of those parameters:  $\langle E_\psi(p,q) \rangle$ . They explore whole classes of solutions at once, and find the values of  $p$  and  $q$  that yield the lowest energy.

What we just described is actually how theorists obtained an approximate solution for the wave function of the ground state of helium.

## Position Operators

Having addressed the average energy of any state  $|\psi\rangle$ , let us now calculate its average position. For simplicity, we will consider only one dimension, and seek the average x-coordinate of state  $|\psi\rangle$ .

Imagine that we have a large collection of atoms in state  $|\psi\rangle$ , and that we measure the x-coordinate of one atom after another. For a very large number of measurements, the average x-coordinate is expected to be:

$$\begin{aligned}\langle x_\psi \rangle &= \int x \text{Prob}(x) dx \\ \langle x_\psi \rangle &= \int \langle \psi | x \rangle x \langle x | \psi \rangle dx \\ \langle x_\psi \rangle &= \int \psi^*(x) x \psi(x) dx\end{aligned}$$

This equation has the same form as the average energy equation, but with  $H$  replaced by  $x$ . If we wish, we can think of  $x$  as an algebraic multiplicative operator:  $x\{\psi(x)\} = x$  times  $\psi(x)$ . We can imagine some state  $|\alpha\rangle$ , such that:

$$\langle x_\psi \rangle = \langle \psi | \alpha \rangle = \int \langle \psi | x \rangle x \langle x | \psi \rangle dx$$

We can expand any  $\langle \psi | \alpha \rangle$  in x-coordinate representation as:

$$\langle \psi | \alpha \rangle = \int \langle \psi | x \rangle \langle x | \alpha \rangle dx$$

Comparing the last two equations, we find:

$$\begin{aligned}x \langle x | \psi \rangle &= \langle x | \alpha \rangle \\ x \psi(x) &= \alpha(x)\end{aligned}$$

Feynman cautions that the first line above does not imply:  $x |\psi\rangle = |\alpha\rangle$ . Cancelling  $\langle x |$  on both sides is equivalent to erroneously cancelling vector  $a$  on both sides of:  $a \bullet b = a \bullet c$ .

In v3p20-9, Feynman says: “We have not bothered to try to get the x-representation of the matrix of [operator  $\underline{x}$ ]. If you are ambitious you can try to show that”:

$$\langle x | \underline{x} | X \rangle = x \delta(x - X)$$

Deriving this relationship is a modest challenge. You can check your answer at the end of this chapter.

If we set  $|X\rangle = |x\rangle$ , we obtain “the amusing result that”:

$$\underline{x} |x\rangle = x |x\rangle$$

This odd equation says the quantum operator  $\underline{x}$  acting on  $|x\rangle$ , a state of definite position  $x$ , has the effect of multiplying that state by  $x$ .

To obtain the average value of  $x^2$ , use the  $x$  operator twice:  $x x$ .

$$\langle x_\psi^2 \rangle = \int \langle \psi | x \rangle x x \langle x | \psi \rangle dx$$

## Average Momentum

Expanding on this approach, we can calculate the average momentum of any state  $|\psi\rangle$ . As before,

$$\langle p_\psi \rangle = \int p \text{Prob}(p) dp$$

Normally, probability equals the magnitude of an amplitude squared, as in  $\text{Prob}(x) = \psi^*(x)\psi(x)$ . However, as discussed in Chapter 22, Feynman chooses a different normalization for momentum states:  $\text{Prob}(p) = \psi^*(p)\psi(p)/h$ . For clarity, I will write  $2\pi\hbar$  instead of  $h$  so you won't have to carefully examine each "h" to see if it has a tiny bar — all h's will be barred.

$$\begin{aligned}\langle p_\psi \rangle &= \int \langle \psi | p \rangle p \langle p | \psi \rangle dp / 2\pi\hbar \\ \langle p_\psi \rangle &= \int \psi^*(p) \mathbf{p} \psi(p) dp / 2\pi\hbar\end{aligned}$$

As above, we can define a new state  $|\beta\rangle$  and a new operator  $\mathbf{p}$ , as follows:

$$|\beta\rangle = \mathbf{p} |\psi\rangle$$

This equation says: quantum operator  $\mathbf{p}$  acting on state  $|\psi\rangle$  produces state  $|\beta\rangle$ . Using  $|\beta\rangle$ , we rewrite the prior equations as:

$$\begin{aligned}\langle p_\psi \rangle &= \int \langle \psi | p \rangle \langle p | \beta \rangle dp / 2\pi\hbar \\ \langle p_\psi \rangle &= \langle \psi | \beta \rangle\end{aligned}$$

With the same logic used for the position operator, we can prove:

$$\langle p | \mathbf{p} | P \rangle = p \delta(p-P)$$

$$\mathbf{p} |p\rangle = p |p\rangle$$

In V320-10, Feynman begins developing a key point. The above equations describe the meaning of the momentum operators  $\mathbf{p}$  and  $\mathbf{p}$  in terms of their action on momentum states  $|p\rangle$  and momentum functions  $\psi(p)$  in a *momentum representation*. Can we also describe their action in a *coordinate representation* — how do they act on position states  $|x\rangle$  and position functions  $\psi(x)$ ?

Let's go back to:

$$\langle p_\psi \rangle = \langle \psi | \beta \rangle$$

and expand this in coordinate representation:

$$\langle p_\psi \rangle = \int \langle \psi | x \rangle \langle x | \beta \rangle dx$$

Fine, but the above equation requires  $\langle x | \beta \rangle = \beta(x)$  in coordinate representation, and all we know so far is  $\langle p | \beta \rangle = \beta(p)$  in momentum representation. Recall from Chapter 22, the following relationship between a state  $|\psi\rangle$  in momentum representation and the same state in coordinate representation.

$$\langle p | \psi \rangle = \int \exp \{ -ipx/\hbar \} \langle x | \psi \rangle dx$$

The above integral is over all values of  $x$ . With this equation, we write two expressions for  $\langle p | \beta \rangle$ :

$$\langle p | \beta \rangle = \int \exp \{ -ipx/\hbar \} \langle x | \beta \rangle dx$$

$$\langle p | \beta \rangle = p \langle p | \psi \rangle = p \int \exp \{ -ipx/\hbar \} \langle x | \psi \rangle dx$$

$$\langle p | \beta \rangle = \int p \exp \{ -ipx/\hbar \} \langle x | \psi \rangle dx$$

Feynman obtains the last line by moving  $p$  inside the integral, saying: “Since the integral is over  $x$  we can put  $p$  inside the integral.” I think this is incorrect... (waiting for lightning to strike)... This  $p$  is not an operator, but the momentum variable that is in general a function of  $x$ . We are most interested in bound states, particles bound by a potential  $V$ . Since  $p^2/2m = E - V$ , and  $V$  is a function of  $x$ ,  $p$  is also a function of  $x$ . Perhaps Feynman intended  $\int dx$  as a “partial” integral, akin to a partial derivative in which only  $x$  varies; however, “partial” integration is beyond ordinary calculus, and merits a thorough explanation. Nonetheless, we continue.

Feynman says that, by comparing the first and last of these three equations, one might be tempted to say:

$$\langle x | \beta \rangle = p \langle x | \psi \rangle = p \psi(x)$$

Feynman says: “No. No!”

Our objective is to find  $\beta(x)$ , but  $p \psi(x)$  is a function of both  $p$  and  $x$ . (I wonder if Feynman was so emphatic because the whole class got it wrong.)

In any case, we must find a way to separate  $p$  from  $\psi(x)$ . The answer is using *integration by parts* and recognizing that:

$$(i\hbar) \partial [\exp \{ -ipx/\hbar \} ] / \partial x = p \exp \{ -ipx/\hbar \}$$

This makes our integral:

$$\langle p | \beta \rangle = \int (i\hbar) \partial [\exp \{ -ipx/\hbar \} ] / \partial x \psi(x) dx$$

Recall the clever technique known as integration by parts:

$$uv = \int d(uv) = \int u dv + \int v du$$

$$\int u dv = uv - \int v du$$

Here we set  $u = i\hbar \psi(x)$ , and  $dv = d[\exp \{ -ipx/\hbar \} ]$ . Integration by parts yields:

$$\langle p | \beta \rangle = i\hbar \psi(x) \exp \{ -ipx/\hbar \} - i\hbar \int \exp \{ -ipx/\hbar \} (\partial \psi / \partial x) dx$$

Here we use the partial derivative because  $\psi$  is a function of multiple variables and we wish to vary

only  $x$ . The above must be evaluated at the integration limits:  $x=-\infty$  and  $x=+\infty$ . For any bound particle,  $\psi(x=\pm\infty)$  must be zero, which eliminates the upper line, leaving us with:

$$\langle p | \beta \rangle = -i\hbar \int \exp\{-ipx/\hbar\} (\partial\psi/\partial x) dx$$

Comparing this with the earlier equation yields:

$$\langle x | \beta \rangle = \beta(x) = -i\hbar \partial\psi/\partial x$$

Going back to our equation for the average momentum in coordinate representation, we can now complete the task.

$$\begin{aligned} \langle p_\psi \rangle &= \int \langle \psi | x \rangle \langle x | \beta \rangle dx \\ \langle p_\psi \rangle &= \int \langle \psi | x \rangle (\hbar/i) (\partial\psi/\partial x) dx \end{aligned}$$

## Operator Correspondence

We now arrive at the key point. In V3p20-11, Feynman emphasizes the following pattern. We compare below amplitude equations with coordinate representation integrals for the average values of energy, position, and momentum.

$$\begin{aligned} \langle E_\psi \rangle &= \langle \psi | \hat{E} \rangle, \text{ with } |\hat{E}\rangle = \hat{H} |\psi\rangle \\ \langle E_\psi \rangle &= \int \psi^*(x) \hat{E}(x) dx, \text{ with } \hat{E}(x) = \hat{H} \psi(x) \end{aligned}$$

$$\begin{aligned} \langle x_\psi \rangle &= \langle \psi | \hat{x} \rangle, \text{ with } |\hat{x}\rangle = \hat{x} |\psi\rangle \\ \langle x_\psi \rangle &= \int \psi^*(x) \hat{x}(x) dx, \text{ with } \hat{x}(x) = x \psi(x) \end{aligned}$$

$$\begin{aligned} \langle p_\psi \rangle &= \langle \psi | \hat{p} \rangle, \text{ with } |\hat{p}\rangle = \hat{p} |\psi\rangle \\ \langle p_\psi \rangle &= \int \psi^*(x) \hat{p}(x) dx, \text{ with } \hat{p}(x) = \hat{P}_x \psi(x) \\ \hat{P}_x &= (\hbar/i) \partial/\partial x \end{aligned}$$

This highlights the connection between the quantum operators and the corresponding algebraic operators, shown below on the left and right, respectively.

$$\begin{aligned} \hat{H} |\psi\rangle &\Leftrightarrow \hat{H} \psi(x); \hat{H} = -(\hbar^2/2m) \partial^2/\partial x^2 + V(x) \\ \hat{x} |\psi\rangle &\Leftrightarrow x \psi(x) = x \psi(x) \\ \hat{p}_x |\psi\rangle &\Leftrightarrow \hat{P}_x \psi(x); \hat{P}_x = (\hbar/i) \partial/\partial x \end{aligned}$$

The extension to three dimensions is evident. The other momentum operators are:

$$\begin{aligned} \hat{p}_y &\Leftrightarrow \hat{P}_y = (\hbar/i) \partial/\partial y \\ \hat{p}_z &\Leftrightarrow \hat{P}_z = (\hbar/i) \partial/\partial z \end{aligned}$$

The three component momentum operators can be combined to form a momentum vector-operator:

$$\mathbf{p} \Leftrightarrow \mathbf{P} = (\hbar/i) (\partial/\partial x, \partial/\partial y, \partial/\partial z)$$

Because both types of operators involve the same physics, most physics books and papers use the same symbols for the quantum operators and the corresponding algebraic operators. We also use the same symbols everywhere in this course except in this chapter where we are focusing on their relationship.

Let's now see how these operators are used.

Note the result of using momentum operators twice.

$$\mathbf{P}_x \mathbf{P}_x = \{(\hbar/i) \partial/\partial x\} \{(\hbar/i) \partial/\partial x\} = -\hbar^2 \partial^2/\partial x^2$$

$$\mathbf{P} \cdot \mathbf{P} = -\hbar^2 \{ \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2 \}$$

$$H = (1/2m) \mathbf{P} \cdot \mathbf{P} + V(\mathbf{r})$$

Feynman says this equation should look familiar “if you haven't forgotten your classical physics”: it says total energy = kinetic energy ( $\mathbf{p} \cdot \mathbf{p}/2m$ ) plus potential energy.

In V3p20-13, Feynman says: “This result has impressed people so much that they try to teach students all about classical physics before quantum mechanics. (We think differently!) But such parallels are often misleading. For one thing, when you have operators, the *order* of various factors is important; but that is not true for the factors in a classical equation.”

In Chapter 23, we defined the momentum operator  $\mathbf{p}_x$  in terms of the displacement operator  $D_x$ . Recall that for an infinitesimal displacement  $dx$  (shifting the  $x$ -axis toward  $+x$  by the amount  $dx$ , or equivalently subtracting  $dx$  from each  $x$ -coordinate value), the action of the displacement operator on state  $|\psi\rangle$  is:

$$D_x(dx) |\psi\rangle = \exp\{ik_x dx\} |\psi\rangle = (1 + ik_x dx) |\psi\rangle$$

$$\mathbf{p}_x = \hbar k_x$$

$$D_x(dx) |\psi\rangle = (1 + i dx \mathbf{p}_x/\hbar) |\psi\rangle$$

$$\partial\psi/\partial x = \{\psi(x+dx) - \psi(x)\} / dx$$

$$\partial\psi/\partial x = \{D_x(-dx) |\psi\rangle - |\psi\rangle\} / dx$$

$$\partial\psi/\partial x = -i (\mathbf{p}_x/\hbar) |\psi\rangle$$

$$\mathbf{p}_x = (\hbar/i) \partial/\partial x$$

This matches what we derived earlier in this chapter. Note we must use  $D_x(-dx)$  to obtain the value of  $\psi$  at  $x+dx$ , due to the definition of  $D_x$ .

Now imagine a collection of particles numbered 1, 2, ...,  $N$ . In one dimension, the wave function is a function of each particle's  $x$ -coordinate:  $\psi(x_1, x_2, \dots, x_N)$ . If we now add  $\delta$  to each  $x$ -coordinate, the wave function becomes:

$$\begin{aligned} & \psi(x_1+\delta, x_2+\delta, \dots, x_N+\delta) - \psi(x_1, x_2, \dots, x_N) \\ &= \delta \{ \partial\psi/\partial x_1 + \partial\psi/\partial x_2 + \dots + \partial\psi/\partial x_N \} \end{aligned}$$

$$\delta (i/\hbar) \mathbf{p}_{sum} \psi$$

$$= \delta (i/\hbar) \{ \mathbf{p}_1 + \mathbf{p}_2 + \dots + \mathbf{p}_N \} \psi$$

$$\mathbf{p}_{sum} = \mathbf{p}_1 + \mathbf{p}_2 + \dots + \mathbf{p}_N$$

Here  $\mathbf{p}_{sum}$  is the algebraic momentum operator for the entire collection of particles and  $\mathbf{p}_j$  is the algebraic momentum operator for the jth particle. We see this result matches the normal rule that the momentum of a total system is the vectorial sum of the momenta of each of its components.

## Angular Momentum Operators

Next let's examine angular momentum operators. In Chapter 23, we defined the angular momentum operator  $\mathbf{J}_z$  in terms of the rotation operator  $R_z$ . We will consider here only orbital angular momentum, and exclude any spin angular momenta. For an infinitesimal rotation by angle  $\epsilon$  (rotating the coordinate system about the z-axis by angle  $\epsilon$ , or equivalently rotating objects or states by angle  $-\epsilon$ ), the action of the rotation operator on state  $|\psi\rangle$  is:

$$R_z(\epsilon) |\psi\rangle = (1 + i \mathbf{J}_z \epsilon/\hbar) |\psi\rangle$$

Figure 31-1 illustrates the relationships between polar coordinates  $(r, \theta, \phi)$  and rectilinear coordinates  $(x, y, z)$  of a point Q, as given by:

$$x = r \sin\theta \cos\phi$$

$$y = r \sin\theta \sin\phi$$

$$z = r \cos\theta$$

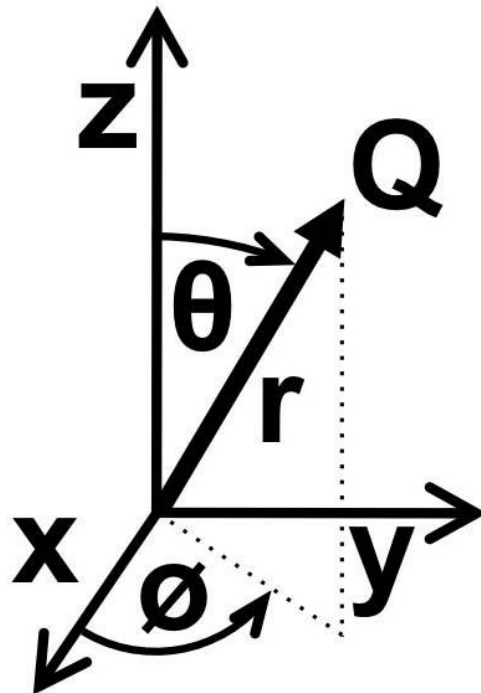


Figure 31-1 Polar & Rectilinear Coordinates

Positive angular momentum in the z-direction,  $J_z > 0$ , corresponds to a steadily increasing azimuthal angle  $\phi$ . To see how the rectilinear coordinates change with this rotation, we take their partial

derivatives with respect to  $\phi$ .

$$\partial x / \partial \phi = -r \sin \theta \sin \phi = -y$$

$$\partial y / \partial \phi = +r \sin \theta \cos \phi = +x$$

$$\partial z / \partial \phi = 0$$

From these derivatives, we can calculate the change in  $|\psi\rangle$  as  $\phi$  increases. To reduce clutter in the next dozen lines, I will use  $\psi$  as an abbreviation of  $|\psi\rangle$ .

$$d\psi = (\partial\psi/\partial x)(\partial x/\partial\phi)d\phi + (\partial\psi/\partial y)(\partial y/\partial\phi)d\phi$$

$$d\psi = \{(\partial\psi/\partial x)(-y) + (\partial\psi/\partial y)(+x)\} d\phi$$

$$d\psi = \{-y \partial/\partial x + x \partial/\partial y\} \psi d\phi$$

We get another equation for the change in  $\psi$  from the rotation operator  $R_z(d\phi)$ :

$$R_z(d\phi) \psi = (1 + \mathbf{J}_z d\phi i/\hbar) \psi$$

$$d\psi = R_z(d\phi)\psi - \psi = \mathbf{J}_z d\phi (i/\hbar) \psi$$

Matching the two equations for  $d\psi$  yields:

$$\mathbf{J}_z d\phi (i/\hbar) \psi = \{-y \partial/\partial x + x \partial/\partial y\} \psi d\phi$$

$$\mathbf{J}_z = (\hbar/i) \{-y \partial/\partial x + x \partial/\partial y\}$$

This can also be written:

$$\mathbf{J}_z = x \mathbf{P}_y - y \mathbf{P}_x$$

The result matches the classical equations for angular momentum:  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$  and  $L_z = xp_y - yp_x$ .

While it is comforting that many quantum operators match classical equations, some do not. Indeed, if all quantum operators matched classical equations, quantum mechanics would not reveal any new physics — what fun would that be?

Thankfully there are differences. Here is one. In classical physics,  $x$  and  $p_x$  commute, as shown by:

$$x p_x - p_x x = 0, \text{ classically}$$

But the corresponding quantum operators do not commute, as we see when the equivalent combination operates on  $\psi(x)$ .

$$\{x \mathbf{P}_x - \mathbf{P}_x x\} \psi(x) = (\hbar/i) \{x \partial/\partial x - \partial/\partial x x\} \psi(x)$$

$$\{x \mathbf{P}_x - \mathbf{P}_x x\} \psi(x) = (\hbar/i) \{x \partial\psi/\partial x - \partial(x\psi)/\partial x\}$$

$$\{x \mathbf{P}_x - \mathbf{P}_x x\} \psi(x) = (\hbar/i) \{x \partial\psi/\partial x - \psi - x \partial\psi/\partial x\}$$

$$\{x \mathbf{P}_x - \mathbf{P}_x x\} \psi(x) = -(\hbar/i) \psi(x)$$

In the second line above,  $\partial/\partial x$  operates on everything to its right, which is the product  $x\psi$ .

While one can do the math on the operators themselves, those still developing their skills will find it easier to analyze the action of these operator combinations on a state or function, such as  $|\psi\rangle$  or  $\psi(x)$ .

The last equation above means the difference of the operators taken in opposite sequences is a non-zero constant, given by:

$$\mathbf{x} \mathbf{P}_x - \mathbf{P}_x \mathbf{x} = i\hbar$$

This demonstrates again that Planck's constant sets the scale at which quantum mechanics diverges from classical physics. On scales much larger than  $\hbar$ , these operators virtually commute and quantum mechanics and classical physics both work. On scales where energy, momentum, time intervals, and distances are comparable to  $\hbar$ , only quantum mechanics correctly describes nature.

The operator combination just discussed has a special name: the *commutator*. For any two algebraic operators or any two quantum operators  $\underline{A}$  and  $\underline{B}$ , their commutator is defined as:

$$[\underline{A}, \underline{B}] = \underline{A} \underline{B} - \underline{B} \underline{A}$$

If  $[\underline{A}, \underline{B}] = 0$ , the operators are said to *commute*; if  $[\underline{A}, \underline{B}]$  is non-zero, they do not commute. This is called the *commutation* rule. Commutators are a staple of quantum mechanics, and also other advanced fields of physics.

One can easily verify that  $y$  and  $p_x$  commute.

$$[y, p_x] \psi = (\hbar/i) \{y \partial \psi / \partial x - \partial (y \psi) / \partial x\} = 0$$

In V3p20-15, Feynman suggests you hone your skills proving:

$$[J_x, J_y] = i\hbar J_z$$

You can check your answer at the end of this chapter.

Feynman points out that non-commuting operators are not unique to quantum mechanics. Matrix products do not commute. Rotation operators do not commute: he suggests you try rotating a book  $90^\circ$  about the  $x$  axis and then  $90^\circ$  about the  $y$ -axis, and compare the result with the same rotations in the opposite order.

## Time Dependent Averages

The final topic on operators in this chapter concerns changes of average values over time. We found earlier that, for any operator  $\underline{A}$ , its average value was determined by:

$$\langle A_\psi \rangle = \langle \psi | \underline{A} | \psi \rangle$$

In what ways can  $\langle A_\psi \rangle$  depend on time? Clearly, a time dependence can only come from either or both quantities on the right side of this equation:  $\psi$  and  $\underline{A}$ . The operator  $\underline{A}$  could, for example, involve

a time-varying potential, such as  $V(x,t)$ . The state  $|\psi\rangle$  could also vary with time; for example, it could be the wave function of a moving particle.

Let's examine the possibilities by taking the time derivative of the above equation.

$$\begin{aligned} d\langle A_\psi \rangle / dt &= (d\langle \psi | / dt) \underline{A} |\psi\rangle \\ &+ \langle \psi | (d\underline{A} / dt) |\psi\rangle \\ &+ \langle \psi | \underline{A} | (d|\psi\rangle / dt) \end{aligned}$$

We can derive expressions for two pieces of this equation. Firstly, the Hamiltonian determines how  $|\psi\rangle$  changes:

$$\underline{H} |\psi\rangle = i\hbar d|\psi\rangle / dt$$

We get a second piece by taking the complex conjugate of the last equation. Recall that  $i^* = -i$ ,  $\langle \psi | = |\psi\rangle^*$ , and  $\underline{H} = \underline{H}^*$ .

$$\langle \psi | \underline{H} = -i\hbar d\langle \psi | / dt$$

Putting these into the time derivative equation yields:

$$\begin{aligned} d\langle A_\psi \rangle / dt &= (\langle \psi | \underline{H} \underline{A} |\psi\rangle (-1/i\hbar) \\ &+ \langle \psi | (d\underline{A} / dt) |\psi\rangle \\ &+ \langle \psi | \underline{A} | \underline{H} |\psi\rangle (+1/i\hbar) \end{aligned}$$

$$\begin{aligned} d\langle A_\psi \rangle / dt &= (i/\hbar) \{ \langle \psi | \underline{H} \underline{A} |\psi\rangle - \langle \psi | \underline{A} \underline{H} |\psi\rangle \} \\ &+ \langle \psi | (d\underline{A} / dt) |\psi\rangle \end{aligned}$$

$$d\langle A_\psi \rangle / dt = \langle \psi | (i/\hbar) [\underline{H}, \underline{A}] + d\underline{A} / dt |\psi\rangle$$

We can define a new operator  $\underline{\alpha}$ , such that:

$$d\langle A_\psi \rangle / dt = \langle \psi | \underline{\alpha} |\psi\rangle$$

We have shown that:

$$\underline{\alpha} = (i/\hbar) [\underline{H}, \underline{A}] + d\underline{A} / dt$$

If operator  $\underline{A}$  has no time-dependence, our result reduces to:

$$\underline{\alpha} = (i/\hbar) [\underline{H}, \underline{A}]$$

Consider an example. Let  $\underline{A}$  be the position operator  $\underline{x}$ , with  $\underline{H}$  given by:

$$\underline{H} = (-\hbar^2/2m) d^2/dx^2 + V(x)$$

$$\underline{\alpha} = d\langle x_\psi \rangle / dt$$

$$\underline{\alpha} |\psi\rangle = (i/\hbar) \{ \underline{H} \underline{x} - \underline{x} \underline{H} \} |\psi\rangle$$

The potential  $V(x)$  cancels in the commutator because  $x$  and  $V$  are both multiplicative operators:  $xV|\psi\rangle = Vx|\psi\rangle$ . We need deal only with the derivative term. (Again, in the next few lines we abbreviate  $|\psi\rangle$  with  $\psi$ .)

$$\begin{aligned} \underline{\alpha} \psi &= (i/\hbar) (-\hbar^2/2m) \{ d^2(x\psi)/dx^2 - x d^2\psi/dx^2 \} \\ \underline{\alpha} \psi &= (\hbar/2im) \{ d/dx (x d\psi/dx + \psi) - x d^2\psi/dx^2 \} \\ &= (\hbar/2im) \{ d\psi/dx + x d^2\psi/dx^2 + d\psi/dx - x d^2\psi/dx^2 \} \\ \underline{\alpha} \psi &= (\hbar/2im) \{ 2d\psi/dx \} \\ \underline{\alpha} \psi &= (\hbar/im) \{ (i/\hbar) \underline{p}_x \} \psi \\ \underline{\alpha} &= d\langle x_\psi \rangle / dt = \underline{p}_x / m \end{aligned}$$

The last equation again matches classical physics:  $dx/dt = p_x/m = v_x$ .

Another example: let  $\underline{A}$  be the position operator  $\underline{p}_x$ .

$$\begin{aligned} \underline{\alpha} &= d\langle \underline{p}_x \rangle / dt \\ \underline{\alpha} |\psi\rangle &= (i/\hbar) \{ \underline{H} \underline{p}_x - \underline{p}_x \underline{H} \} |\psi\rangle \\ \underline{\alpha} |\psi\rangle &= \{ \underline{H} d/dx - d/dx \underline{H} \} |\psi\rangle \end{aligned}$$

The derivatives  $d/dx$  and  $d^2/dx^2$  commute, leaving only  $V(x)$  in the commutator.

$$\begin{aligned} \underline{\alpha} |\psi\rangle &= \{ V d/dx - d/dx V \} |\psi\rangle \\ \underline{\alpha} |\psi\rangle &= V d|\psi\rangle/dx - V d|\psi\rangle/dx - |\psi\rangle dV/dx \\ \underline{\alpha} &= d\langle \underline{p}_x \rangle / dt = -dV/dx \end{aligned}$$

This also matches classical physics:  $dp/dt = ma = F = -dV/dx$ .

Feynman stresses that we have examined *average* values produced by quantum and algebraic operators acting on large numbers of entities. These equations tell us nothing about what happens at the individual particle level.

The key distinguishing feature of quantum mechanics that we have identified is that  $xp - px$  is not zero as in classical physics, but rather it equals some small number  $i\hbar$ . Small yes, but as Feynman says: “the whole wondrous complications of interference, waves, and all, result from the little fact that” this small number is not zero.

In V3p20-17, Feynman ends this lecture saying:

“The history of this idea is also interesting. Within a period of a few months in 1926, Heisenberg and Schrödinger independently found correct laws to describe atomic mechanics. Schrödinger invented his wave function  $\psi(x)$  and found his equation. Heisenberg, on the other hand, found that nature could be described by classical equations, except that  $xp - px$  should be equal to  $i\hbar$ , which he could make happen by defining them in terms of special kinds of matrices. In our language he was using the energy-representation, with its matrices. Both Heisenberg’s matrix

algebra and Schrödinger's differential equation explained the hydrogen atom. A few months later Schrödinger was able to show that the two theories were equivalent—as we have seen here. But the two different mathematical forms of quantum mechanics were discovered independently.”

## Problem: Position Operator Matrix

Feynman suggested the challenge of deriving the matrix elements of the position operator.

Recall the definition of the Dirac delta function from Chapter 22:

$$\langle x|\psi\rangle = \int \delta(x-X) \langle X|\psi\rangle dX$$

Earlier in this chapter, we found:

$$\langle x_\psi\rangle = \int \langle \psi|x\rangle x \langle x|\psi\rangle dx$$

Now replace  $\langle x|\psi\rangle$  using the Dirac delta function.

$$\langle x_\psi\rangle = \int \langle \psi|x\rangle x \left\{ \int \delta(x-X) \langle X|\psi\rangle dX \right\} dx$$

Since the delta function is zero everywhere except  $x=X$ , we can rearrange the double integral to yield:

$$\langle x_\psi\rangle = \iint \langle \psi|x\rangle x \delta(x-X) \langle X|\psi\rangle dX dx$$

Now compare this with the equation derived earlier for the matrix elements of the quantum Hamiltonian operator  $\underline{H}$  in a continuous basis:

$$\langle E_\psi\rangle = \iint \langle \psi|x\rangle \langle x|\underline{H}|X\rangle \langle X|\psi\rangle dx dX$$

By analogy, the matrix elements of the quantum position operator  $x$  are given by:

$$\langle x_\psi\rangle = \iint \langle \psi|x\rangle \langle x|\underline{x}|X\rangle \langle X|\psi\rangle dx dX$$

This means:

$$\langle x|\underline{x}|X\rangle = x \delta(x-X)$$

QED.

## Problem: $J_x, J_y$ Commutator

Feynman suggested the challenge of calculating  $[J_x, J_y]$ . Here is the answer.

$$[J_x, J_y] \psi = J_x J_y \psi - J_y J_x \psi$$

$$\begin{aligned} J_x J_y \psi &= (\hbar/i)^2 (y\partial/\partial z - z\partial/\partial y) (z\partial/\partial x - x\partial/\partial z) \psi \\ J_x J_y \psi &= (\hbar/i)^2 (y\partial/\partial z - z\partial/\partial y) (z\partial\psi/\partial x - x\partial\psi/\partial z) \\ J_x J_y \psi &= (\hbar/i)^2 (y\partial\psi/\partial x + yz\partial^2\psi/\partial z\partial x \\ &\quad -yx\partial^2\psi/\partial z^2 - z^2\partial^2\psi/\partial y\partial x + zx\partial^2\psi/\partial y\partial z) \end{aligned}$$

$$\begin{aligned} J_y J_x \psi &= (\hbar/i)^2 (z\partial/\partial x - x\partial/\partial z) (y\partial/\partial z - z\partial/\partial y) \psi \\ J_y J_x \psi &= (\hbar/i)^2 (z\partial/\partial x - x\partial/\partial z) (y\partial\psi/\partial z - z\partial\psi/\partial y) \\ J_y J_x \psi &= (\hbar/i)^2 (zy\partial^2\psi/\partial x\partial z - z^2\partial^2\psi/\partial x\partial y \\ &\quad -xy\partial^2\psi/\partial z^2 + x\partial\psi/\partial y + xz\partial^2\psi/\partial z\partial y) \end{aligned}$$

Partial derivatives commute with one another:  $\partial^2\psi/\partial z\partial y = \partial^2\psi/\partial y\partial z$ , etc. Also the position operators commute because they are simply multipliers. This means all the above terms with second order partial derivatives cancel when we subtract  $J_y J_x \psi$  from  $J_x J_y \psi$  — that's four cancelling pairs. The only parts that do not commute have the form  $[x, \partial/\partial x]\psi$ , where one partial derivative is eliminated. After all the cancellations, only two first order partial derivative terms remain.

$$\begin{aligned} [J_x, J_y] \psi &= (\hbar/i)^2 (y \partial\psi/\partial x - x \partial\psi/\partial y) \psi \\ [J_x, J_y] \psi &= (\hbar/i)^2 (-J_z)(i/h) \psi \\ [J_x, J_y] &= i\hbar J_z \end{aligned}$$

QED.

## Chapter 31 Review: Key Ideas

- This chapter illuminates different conventions used in the physics literature. **Quantum operators** act on quantum states producing new states. By contrast, **algebraic operators** act on functions producing new functions. The distinction is somewhat pedantic. In this chapter only, we underline for clarity quantum operators and use bold italics for algebraic operators. Standard mathematical operators such as  $d/dx$  are not changed. For example:

Quantum vs. Algebraic

$$\begin{aligned} \underline{H} |\psi\rangle &\Leftrightarrow \underline{H} \psi(x); \underline{H} = -(\hbar^2/2m)\partial^2/\partial x^2 + V(x) \\ \underline{x} |\psi\rangle &\Leftrightarrow \underline{x} \psi(x) = x \psi(x) \\ \underline{p}_x |\psi\rangle &\Leftrightarrow \underline{P}_x \psi(x); \underline{P}_x = (\hbar/i) \partial/\partial x \end{aligned}$$

Operators  $\underline{P}_y$  and  $\underline{P}_z$  are similarly defined, and with  $\underline{P}_x$  form a vector-operator that condenses the Hamiltonian equation to:

$$\underline{H} = (1/2m) \underline{P} \cdot \underline{P} + V(\underline{r})$$

For any operator  $\underline{A}$ , its *Hermitian adjoint*  $\underline{Z}$  is defined by:

$$\underline{Z}_{ij} = (\underline{A}_{ji})^*$$

If  $\underline{Z} = \underline{A}$ ,  $\underline{A}$  is called *Hermitian* or *self-adjoint*.

- If  $M$  independent **measurements** are made on quantum entities in a uniform population (such as atoms in the same state), each measured value must equal one of the eigenvalues of that population. Assume a quantum system with discrete eigenstates  $|j\rangle$ ,  $j=0, 1, \dots$  and eigenvalues  $E_j$ . Let the number of measurements with value  $E_j$  be  $M_j$ , with  $M=\sum_j M_j$ . The average measured value is:

$$\langle E_\psi \rangle = \sum_j (M_j/M) E_j$$

For large  $M$ , each  $M_j/M$  approaches the *expectation value*  $\text{Prob}(j)$ , the probability of  $|j\rangle$ , and  $\sigma_{mj}$ , the root-mean-square deviation of  $M_j$ , equals:

$$\sigma_{mj} = \sqrt{\{ M p(1-p) \}}, \text{ with } p=\text{Prob}(j)$$

If the probability of any particular outcome is small ( $p \ll 1$ ), this becomes:

$$\sigma_{mj} = \sqrt{\{ M_j \}}$$

For any physically observable quantity  $A$  associated with a quantum operator  $\underline{A}$ , the average value of  $A$  for any state  $|\psi\rangle$  can be expressed independently of basis states:

$$\langle A_\psi \rangle = \langle \psi | \underline{A} | \psi \rangle$$

For a system with continuous, normalized states, this becomes:

$$\langle A_\psi \rangle = \iint \psi^*(x) \langle x | \underline{A} | X \rangle \psi(X) dx dX$$

- **Commutators** have important consequences in quantum mechanics. For operators  $\underline{A}$  and  $\underline{B}$ , their commutator is:

$$[\underline{A}, \underline{B}] = \underline{A} \underline{B} - \underline{B} \underline{A}$$

If  $[\underline{A}, \underline{B}] = 0$ , the operators *commute*, otherwise they do not commute. Examples are:

$$\underline{x} \underline{P}_x - \underline{P}_x \underline{x} = i\hbar$$

$$\underline{x} \underline{P}_y - \underline{P}_y \underline{x} = 0$$

$$[\underline{J}_x, \underline{J}_y] = i\hbar \underline{J}_z$$

This demonstrates that  $\hbar$  sets the scale at which quantum mechanics diverges from classical physics. On scales  $\gg \hbar$ , these operators virtually commute and quantum mechanics and classical physics both work. On scales with energy, momentum, time intervals, and distances comparable to  $\hbar$ , only quantum mechanics correctly describes nature.

- The **time evolution** of the average  $\langle A_\psi \rangle$  of a large sample of measurements corresponding to operator  $A$  is:

$$d\langle A_{\psi} \rangle / dt = \langle \psi | (i/\hbar) [\underline{H}, \underline{A}] + d\underline{A}/dt | \psi \rangle$$

Examples and their corresponding classical analogs are:

$$d\langle x_{\psi} \rangle / dt = \langle p_x \rangle / m \Leftrightarrow dx/dt = p/m = v$$

$$d\langle p_x \rangle / dt = -dV/dx \Leftrightarrow dp/dt = ma = -dV/dx = F$$

# Chapter 32

## Probability & Wave Functions

---

Feynman's last lecture contains profound insights into quantum mechanics in general, and an extensive exploration of specific low-temperature phenomena. I have rearranged this material, putting the general insights in this chapter and the low-temperature physics in the next chapter.

### The Continuity of Probability

A fundamental principle of quantum mechanics is that the probability,  $\text{Prob}(\mathbf{r}, t)$ , of finding a particle at position  $\mathbf{r}$  at time  $t$  equals  $|\psi(\mathbf{r}, t)|^2$ , the square of the magnitude of its wave function at  $\mathbf{r}$  and  $t$ . The total probability of finding a particle anywhere must always equal 1.

In V3p21-3, Feynman emphasizes that total probability is not only conserved globally, but more importantly, it is also conserved locally. Let's understand what this means by first considering a simpler case: the global and local conservation of electric charge.

Global charge conservation means the net amount of electric charge everywhere is the same at every time  $t$ . By "net amount" we mean the total of all positive charges minus the total of all negative charges. Local charge conservation adds an additional requirement: if net charge decreases in one region, charge must *flow* out through that region's boundaries into neighboring regions. This means electric charge cannot disappear here and reappear at some distant location; it must move continuously from one region to another, being conserved at every step along the way. Local conservation tells us something important about *how* charge is conserved globally, and inspires us to understand electric *currents* of flowing charge.

In this section, Feynman suggests an analogy between conservation of charge and conservation of probability. Global probability conservation means the integral of  $\text{Prob}(\mathbf{r}, t)$  over all  $\mathbf{r}$  equals 1 at every time  $t$ . Feynman suggests local probability conservation means that, if total probability decreases in one region, probability must "flow" out through that region's boundaries into neighboring regions.

Can we find a probability "current" associated with this "flow"? Feynman says: "This current would be a vector which could be interpreted this way—the x-component would be the net probability per second and per unit area that a particle passes in the x-direction across a plane parallel to the yz-plane. Passage toward +x is considered a positive flow, and passage in the opposite direction, a negative flow."

We begin with the equation for probability, and take its partial derivative with respect to time.

$$\begin{aligned}\text{Prob}(\mathbf{r},t) &= \psi^*(\mathbf{r},t) \psi(\mathbf{r},t) \\ \partial(\text{Prob})/\partial t &= \psi^* \partial\psi/\partial t + \psi \partial\psi^*/\partial t\end{aligned}$$

We next use the Hamiltonian to find  $\partial\psi/\partial t$ .

$$\begin{aligned}i\hbar \partial\psi/\partial t &= H \psi \\ \partial\psi/\partial t &= (-i/\hbar) H \psi \\ \partial\psi^*/\partial t &= (+i/\hbar) H^* \psi^* \\ \partial(\text{Prob})/\partial t &= (-i/\hbar) \psi^* H \psi + (+i/\hbar) \psi H^* \psi^* \\ i\hbar \partial(\text{Prob})/\partial t &= \psi^* H \psi - \psi H^* \psi^*\end{aligned}$$

We now use the normal Hamiltonian, in operator form.

$$H = (1/2m) \mathbf{P} \bullet \mathbf{P} + V(\mathbf{x})$$

$$\begin{aligned}i\hbar \partial(\text{Prob})/\partial t &= \psi^* V \psi - \psi V^* \psi^* \\ &\quad + (1/2m) \{ \psi^* \mathbf{P} \bullet \mathbf{P} \psi - \psi \mathbf{P} \bullet \mathbf{P} \psi^* \}\end{aligned}$$

The two potential terms in the upper line cancel. For the lower line, Feynman provides yet another clever trick. Consider the following, shown in one-dimension to reduce clutter, and recall that  $P_x = (\hbar/i)\partial/\partial x$ :

$$\begin{aligned}\partial/\partial x \{ \psi^* P_x \psi \} &= \partial/\partial x \{ \psi^* (\hbar/i) \partial\psi/\partial x \} \\ &= (\hbar/i) \psi^* \partial^2\psi/\partial x^2 + (\hbar/i) \partial\psi^*/\partial x \partial\psi/\partial x \\ \partial/\partial x \{ \psi^* P_x \psi \}^* &= \partial/\partial x \{ \psi (-\hbar/i) \partial\psi^*/\partial x \} \\ &= -(\hbar/i) \psi \partial^2\psi^*/\partial x^2 - (\hbar/i) \partial\psi/\partial x \partial\psi^*/\partial x\end{aligned}$$

Now add the last two equations.

$$\begin{aligned}\partial/\partial x \{ (\psi^* P_x \psi) + (\psi^* P_x \psi)^* \} \\ = (\hbar/i) \psi^* \partial^2\psi/\partial x^2 - (\hbar/i) \psi \partial^2\psi^*/\partial x^2\end{aligned}$$

Restoring y- and z-derivatives is straightforward:

$$\begin{aligned}(i/\hbar) \mathbf{P} \bullet \{ (\psi^* \mathbf{P} \psi) + (\psi^* \mathbf{P} \psi)^* \} \\ = (\hbar/i) (i/\hbar)^2 \{ \psi^* \mathbf{P} \bullet \mathbf{P} \psi - \psi \mathbf{P} \bullet \mathbf{P} \psi^* \}\end{aligned}$$

$$\begin{aligned}\mathbf{P} \bullet \{ (\psi^* \mathbf{P} \psi) + (\psi^* \mathbf{P} \psi)^* \} \\ = \{ \psi^* \mathbf{P} \bullet \mathbf{P} \psi - \psi \mathbf{P} \bullet \mathbf{P} \psi^* \}\end{aligned}$$

We can now rewrite our primary equation as:

$$i\hbar \partial(\text{Prob})/\partial t = +(1/2m) \mathbf{P} \bullet \{ (\psi^* \mathbf{P} \psi) + (\psi^* \mathbf{P} \psi)^* \}$$

$$\partial(\text{Prob})/\partial t = -(i/\hbar) \mathbf{P} \bullet \{(\psi^* \mathbf{P} \psi) + (\psi^* \mathbf{P} \psi)^*\} / 2m$$

Note:  $(i/\hbar) \mathbf{P}$  is the divergence vector-operator  $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$ , written as an inverted  $\Delta$ . It is a real (not a complex) quantity.

Feynman notes that the expression in  $\{ \}$ 's is real, being the sum of a quantity and its complex conjugate. This should not be surprising since it comes from the time derivative of  $\psi^* \psi$ , a quantity that is always real.

We can now define a probability current  $\hat{\mathbf{J}}$ , and simplify the probability derivative.

$$\hat{\mathbf{J}} = \{(\psi^* \mathbf{P} \psi) + (\psi^* \mathbf{P} \psi)^*\} / 2m$$

$$\hat{\mathbf{J}} = \{(\psi^* \mathbf{P} \psi) + (\psi \mathbf{P}^* \psi^*)\} / 2m$$

$$\partial(\text{Prob})/\partial t = -(i/\hbar) \mathbf{P} \bullet \hat{\mathbf{J}}$$

Or if you prefer:

$$\partial(\text{Prob})/\partial t + (i/\hbar) \mathbf{P} \bullet \hat{\mathbf{J}} = 0$$

This equation says the time derivative of probability plus the divergence of the probability current equals zero, within whatever volume we choose. With Gauss's divergence theorem, this means the change in probability within any volume equals the probability current flowing into that volume (inward flow is a negative divergence). This means probability can change only by moving continuously from one place to another: probability is conserved locally.

Consider an example of a particle moving from region A to a neighboring region B. The probability in region A decreases and its probability current has a positive divergence. At the same time, the probability in region B increases and its probability current has a negative divergence. Hence probability flows out of region A into region B, just as we would expect.

We will come back to the probability current in the next chapter.

## The Meaning of Wave Functions

In V3p21-6, Feynman says:

“When Schrödinger first discovered his equation he discovered the conservation law [of probability] as a consequence of his equation. But he imagined incorrectly that  $\psi^* \psi$  was the electric charge density of the electron and that  $\hat{\mathbf{J}}$  was the electric current density, so he thought that the electrons interacted with the electromagnetic field through these charges and currents. When he solved his equations for the hydrogen atom and calculated  $\psi$ , he wasn't calculating the probability of anything—there were no amplitudes [in quantum theory] at that time—the interpretation was completely different. The atomic nucleus was stationary but there were currents moving around; the charges  $\psi^* \psi$  and currents  $\hat{\mathbf{J}}$  would generate electromagnetic fields and the thing would radiate light. He soon found on doing a number of problems that it didn't

work out quite right. It was at this point that [Max] Born made an essential contribution to our ideas regarding quantum mechanics. It was Born who correctly (as far as we know) interpreted the  $\psi$  of the Schrödinger equation in terms of a probability amplitude—that very difficult idea that the square of the amplitude is not the charge density but is only the probability per unit volume of finding an electron there, and that when you do find the electron some place the entire charge is there. That whole idea is due to Born.

“The wave function  $\psi(\mathbf{r})$  for an electron in an atom does not, then, describe a smeared-out electron with a smooth charge density. The electron [will be found] either here, or there, or somewhere else, but wherever it is, it is a point charge. On the other hand, think of a situation in which there are an enormous number of particles in exactly the same state, a very large number of them with exactly the same wave function. Then what? One of them is here and one of them is there, and the probability of finding any one of them at a given place is proportional to  $\psi^*\psi$ . But since there are so many particles, if I look in any volume  $dxdydz$  I will generally find a number close to  $\psi^*\psi dxdydz$ . So in a situation in which  $\psi$  is the wave function for each of an enormous number of particles which are all in the same state,  $\psi^*\psi$  *can* be interpreted as the density of particles.

“So in the situation in which we can have very many particles in exactly the same state, ...the charge density and the electric current can be calculated directly from the wave functions and the wave functions take on a physical meaning which extends into classical, macroscopic situations.

“Something similar can happen with neutral particles. When we have the wave function of a single photon, it is the amplitude to find a photon somewhere. Although we haven’t ever written it down there is an equation for the photon wave function analogous to the Schrödinger equation for the electron. The photon equation is just the same as Maxwell’s equations for the electromagnetic field, and the wave function is the same as the vector potential  $A$ . The wave function turns out to be just the vector potential. ... The moment that you have billions in the same state (that is, in the same electromagnetic wave), you can measure the wave function, which is the vector potential, directly.

“Now the trouble with the electron [a fermion] is that you cannot put more than one in the same state. Therefore, it was long believed that the wave function of the Schrödinger equation would never have a macroscopic representation analogous to the macroscopic representation of the amplitude for photons... it is now realized that the phenomena of superconductivity presents us with just this situation.”

We explore superconductivity in the next chapter.

## Chapter 32 Review: Key Ideas

- The **probability**  $\text{Prob}(\mathbf{r},t)$  of finding a particle at position  $\mathbf{r}$  at time  $t$  equals  $|\psi(\mathbf{r},t)|^2$ , the square of the magnitude of its wave function at  $\mathbf{r}$  and  $t$ . The total probability of finding a particle anywhere must always equal 1. This means probability is conserved globally. But more importantly, probability is

conserved locally as well. This is analogous to the conservation of electric charge.

We define the **probability current**  $\hat{\mathbf{J}}$ , and its relationship to the change of probability.

$$\hat{\mathbf{J}} = \{(\psi^* \mathbf{P} \psi) + (\psi^* \mathbf{P} \psi)^*\} / 2m$$

$$\partial(\text{Prob})/\partial t + (\mathbf{i}/\hbar) \mathbf{P} \cdot \hat{\mathbf{J}} = 0$$

The time derivative of probability plus the divergence of probability current equals zero, integrated over any volume we choose. With Gauss's divergence theorem, this means the change in probability within any volume equals the probability current flowing into that volume (inward flow is a negative divergence). This means probability is conserved locally, within any selected volume.

- The wave function describes the amplitude of finding the electron at each location, at each moment in time. But this does not mean an electron is smeared-out with a smooth charge density. If searched for, an electron will be found at a specific point, with all of its charge at that point.

If an immense number of particles exist in exactly the same state, Feynman says: “ $\psi^* \psi$  *can* be interpreted as the density of particles.”

- The wave function for photons is the same as the electromagnetic vector potential  $A$ , which we can directly measure.

The remarkable thing about low-temperature phenomena, such as superconductivity, is that they allow an immense number of fermions to combine into the same state, allowing direct measurement of their wave function. The manner in which pairs of fermions can behave like bosons is the subject of the next chapter.

# Chapter 33

## Superconductivity: QM Writ Large

---

The last of the *Feynman Lectures on Physics* begins with:

“This lecture is only for entertainment. I would like to give the lecture in a somewhat different style—just to see how it works out. It’s not a part of the course—in the sense that it is not supposed to be a last minute effort to teach you something new. But, rather, I imagine that I’m giving a seminar ... to a more advanced audience... So in this lecture I’ll describe the ideas all the way along but just give you the *results* of the computations. You should realize that you’re not supposed to understand everything immediately...”

Let me comment on what Feynman means by “seminar.” A physics seminar is not a normal course for physics students; the audience consists of professors, Ph.D. physicists, and graduate students. Seminars are primarily venues for presenting frontier research to colleagues engaged in the same or in closely related fields. Presenters generally assume that the audience knows everything that has already been discovered on the topic, and building on that, they describe their efforts to discover new science. Seminars are an opportunity for presenters to impress their colleagues and further their careers. Seminars are auditions, not classes. Attendees who are world-class experts in the particular topic will likely understand every word and will learn something new. Others hope to understand some of the discussion and learn what they can.

When Feynman gave seminars at Caltech, almost every physics faculty member and graduate student attended. Legend has it that at one seminar, Feynman dashed into the room, by-passed any introductory pleasantries, and immediately began furiously writing on the blackboard. One hour and many blackboards later, Feynman turned to the audience for the first time and said: “Was that clear, Murray?” Apparently, he hoped one person had understood — the other Nobel Laureate in the room, Murray Gell-Mann.

In this “seminar”, Feynman continues:

“My subject is the Schrödinger equation in a classical setting—the case of superconductivity.

“Ordinarily, the wave function which appears in the Schrödinger equation applies to only one or two particles. And the wave function itself is not something that has a classical meaning—unlike the electric field... Nevertheless, there are some situations in which a quantum mechanical wave function *does* have classical significance...”

“It is the purpose of this lecture to show a connection between quantum mechanics and large-

scale effects...a special situation in which quantum mechanics will produce its own characteristic effects on a large or ‘macroscopic’ scale.”

Superconductivity is the property of conducting electric current without resistance — not with a very small resistance, but with *absolutely zero* resistance. Many materials, but not all, are superconducting at sufficiently low temperatures.

Superconductivity was discovered in 1911 by Dutch physicist Heike Kamerlingh Onnes, who received the Nobel Prize in 1913.

The essential distinction between quantum and classical physics is discrete versus continuous states. In classical physics, particle energies are continuous, they can have any value across a wide range, whereas quantum mechanics restricts particle energies to a limited, discrete set of values.

As Feynman explains in V3p21-1, at extreme low temperatures higher-energy atomic states are “frozen out”, and behavior becomes dominated by the quantum mechanical discreteness of the remaining states. This is why quantum effects can produce the macroscopic phenomenon of superconductivity.

Boltzmann’s law says the population of a state with energy  $E$  is proportional to  $\exp\{-E/kT\}$  (*Feynman Simplified 1B* Chapter 16).

When  $kT$  is large, many states are populated with atoms and particles jumping freely from one discrete energy level to another. The properties of different states are averaged, resulting in nearly continuous behavior that is adequately described by classical physics.

When  $kT$  is small, energy level differences become formidable barriers. Only a few low-energy states are populated. Here, Feynman says: “the quantum mechanical character of that ground state” is apparent even macroscopically.

## Schrödinger’s Equation in Magnetic Field

We will shortly develop the Schrödinger equation for an electron in a magnetic field. To obtain this, we will employ several concepts that Feynman presents in his Volume II. Let’s briefly review those.

The first of these concerns *vector fields*. At each point  $\mathbf{r}$ , a normal function  $f(\mathbf{r})$  has a value that is simply a number. An example is  $T(\mathbf{r})$ , the temperature throughout Earth’s atmosphere. Such a function is sometimes called a *scalar field*, with “field” indicating an array of quantities defined throughout space and “scalar” identifying these quantities as simple numbers. A *vector field* is a function that at each point  $\mathbf{r}$  has a value that is a vector. An example is  $\mathbf{W}(\mathbf{r})$ , the wind velocity vector field; at each point  $\mathbf{r}$  throughout Earth’s atmosphere, the wind has a speed and direction represented by a three-component velocity vector.

The second concept is that a magnetic field  $\mathbf{B}$  is described in terms of a *vector potential*  $\mathbf{A}$ , according

to:

$$\mathbf{B} = \text{curl } \mathbf{A}$$

The above vector equation can be written in component notation, without using the *curl* operator, as:

$$B_x = \partial A_z / \partial y - \partial A_y / \partial z$$

$$B_y = \partial A_x / \partial z - \partial A_z / \partial x$$

$$B_z = \partial A_y / \partial x - \partial A_x / \partial y$$

The term *vector potential* is an extension of *vector field*:  $\mathbf{A}$ , is a vector field that specifies an electrodynamic potential complimenting the electrostatic potential due to electric charges. Vector potential  $\mathbf{A}$ , is associated with currents, moving electric charges.

The third concept from Volume II describes how vector potential  $\mathbf{A}$  affects moving charges. In V3p21-1, Feynman says:

“The principle that describes the behavior of quantum mechanics in a vector potential is very simple. The amplitude that a particle goes from one place to another along a certain route when there’s a field present is the same as the amplitude that it would go along the same route when there’s no field, multiplied by the exponential of the line integral of the vector potential, times the electric charge divided by Planck’s constant.” (See V2p15-9.)

While all that may be “very simple” to Feynman, let’s go through this carefully. We write his statement mathematically as:

$$\langle F|S \rangle_{\text{in } \mathbf{A}} = \langle F|S \rangle_{\mathbf{A}=0} \exp \left\{ i(q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s} \right\}$$

Here, the integral is along some path that we label  $\Gamma$ , from start S to finish F. This is what Feynman called the “route.” The path can have any shape whatsoever. At each point P along path  $\Gamma$ , the integrand is  $\mathbf{A} \cdot d\mathbf{s}$ , where  $d\mathbf{s}$  is the tangent vector at that point, as shown in Figure 33-1.

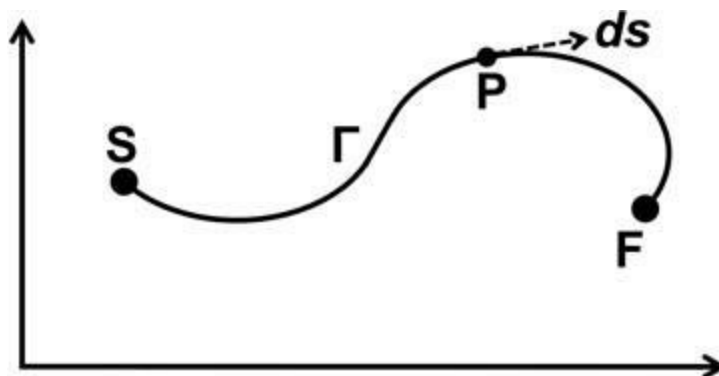


Figure 33-1 Path Integral Along  $\Gamma$  from S to F

This type of integral is appropriately called a *path integral*. Each path integral requires specifying not only the end points S and F, but also the entire path. In general, a path integral has a different value for different paths, even if the end points are the same.

In this case, the path integral is:

$$\int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}$$

The presence of a magnetic field, specified by  $\mathbf{A}$ , modifies the phase of a particle's amplitude as it traverses path  $\Gamma$ . We express this phase change by multiplying the particle's amplitude by the following complex exponential:

$$\exp\{ i(q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s} \}$$

With no magnetic field present ( $\mathbf{A}=0$ ), we denote the amplitude for a particle to go from S to F as:

$$\langle F|S \rangle_{\mathbf{A}=0}$$

With a magnetic field, Feynman says this amplitude is modified to:

$$\langle F|S \rangle_{\text{in } \mathbf{A}} = \langle F|S \rangle_{\mathbf{A}=0} \exp\{ i(q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s} \}$$

Putting this modified amplitude into the Schrödinger equation gives us the equation for a particle with charge  $q$  in a magnetic field specified by the vector potential  $\mathbf{A}$ . For non-relativistic velocities and ignoring spin, Feynman says this equation is:

$$i\hbar \partial\psi/\partial t = (1/2m) (\mathbf{P}-q\mathbf{A}) \cdot (\mathbf{P}-q\mathbf{A}) \psi + V\psi$$

Here,  $V$  is the electrostatic potential energy as before, and the normal momentum operator  $\mathbf{P}$  is augmented by  $-q$  times the vector potential  $\mathbf{A}$ .

Feynman provides a justification of this equation. The rather lengthy mathematics is provided at the end of this chapter.

Recall our prior discussion regarding probability currents. Without a magnetic field, we found

$$\hat{\mathbf{J}} = \{ (\psi^* \mathbf{P} \psi) + (\psi \mathbf{P}^* \psi^*) \} / 2m$$

$$\partial(\text{Prob})/\partial t = -(i/\hbar) \mathbf{P} \cdot \hat{\mathbf{J}}$$

If we add a magnetic field, the only change is replacing  $\mathbf{P}$  with  $\mathbf{P}-q\mathbf{A}$ .

## Two Kinds of Momentum

We are quite accustomed to momentum being mass times velocity:  $\mathbf{p} = m\mathbf{v}$ . Yet in a magnetic field, we find that both the probability current and the Schrödinger equation employ a different quantity in the role of momentum, one that involves the vector potential  $\mathbf{A}$ . What does this mean?

Feynman says both types of momentum are important, in different roles. He defines these as:

“mv-momentum”:  $\mathbf{p} = m\mathbf{v}$

“p-momentum”:  $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$

He adds that the same two types of momentum arise in classical physics as well. This is well described in a classic physics textbook titled *Classical Electrodynamics* by J. D. Jackson. That fine book is *not* easier to read than the *Feynman Lectures*.

In V3p21-5, Feynman explains why we need two types of momentum. Consider a solenoid that initially has no current flowing through it. Imagine that at time  $t=0$ , a current  $J$  begins to flow, growing rapidly and creating both a magnetic field  $\mathbf{B}$  and an electric field  $\mathbf{E}$ , as illustrated in Figure 33-2.

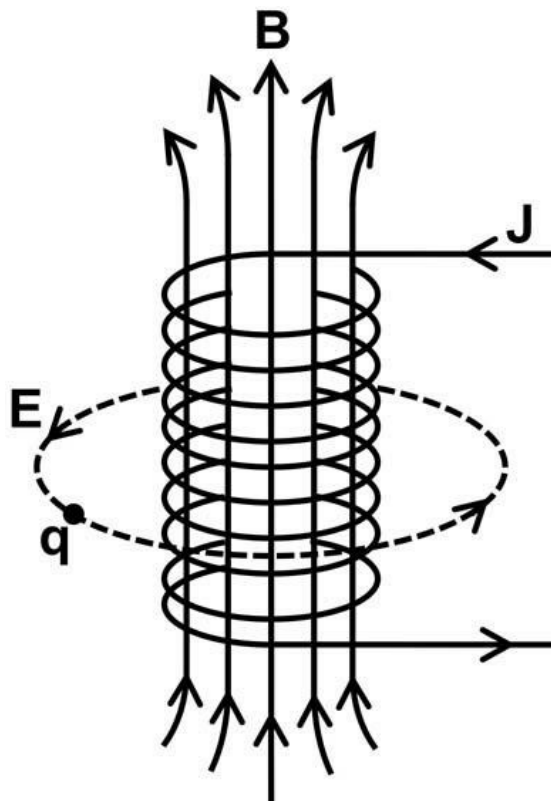


Figure 33-2 Charge  $q$  Near Solenoid

A particle with mass  $m$  and charge  $q$  is initially at rest, lying near to but outside the solenoid windings.

The vector potential  $\mathbf{A}$  is zero at  $t=0$ , but grows rapidly as the current increases. At time  $t=\Delta t$ , the current and the vector potential both attain their maximum values:  $J_{\max}$  and  $\mathbf{A}_{\max}$  respectively. Maxwell's equations say: (1) the path integral of  $\mathbf{A}$  around the solenoid (around the dashed circle for example) equals the flux of  $\mathbf{B}$  through that circle; and (2) a changing  $\mathbf{A}$  creates an electric field  $\mathbf{E}$ , in accordance with:

$$\mathbf{E} = -d\mathbf{A}/dt$$

During the brief time interval  $\Delta t$ , the electric field exerts a force  $\mathbf{F}$  on charge  $q$ , changing the particle's momentum  $\mathbf{p}$ . These relationships apply:

$$\mathbf{F} = q\mathbf{E}$$

$$\begin{aligned} d\mathbf{p}/dt &= \mathbf{F} = -q \, d\mathbf{A}/dt \\ \Delta\mathbf{p} &= \int \mathbf{F} \, dt = -q \int (d\mathbf{A}/dt) \, dt \\ \Delta\mathbf{p} &= -q \, \mathbf{A}_{max} \end{aligned}$$

If  $J$ , and therefore  $\mathbf{A}$ , increase linearly during  $\Delta t$ , the rate of change of  $\mathbf{p}$  during  $\Delta t$  is:

$$d\mathbf{p}/dt = -q \, \mathbf{A}_{max} / \Delta t$$

If  $\Delta t$  is infinitesimal, the particle's momentum change  $\Delta\mathbf{p}$  occurs almost instantaneously. But, the particle's wave function  $\psi$  does not change as rapidly. The rate of change of  $\psi$  is determined by the Hamiltonian equation. Since there is no electrostatic potential  $V$  in this case, the Hamiltonian is:

$$i\hbar 2m \, \partial\psi/\partial t = (\mathbf{P}-q\mathbf{A}) \cdot (\mathbf{P}-q\mathbf{A}) \, \psi$$

We see that  $\partial\psi/\partial t$  is initially zero and increases in proportion to  $\mathbf{A}^2$ , which can be much less than  $\mathbf{A}_{max}/\Delta t$  if  $\Delta t$  is small enough. While the particle's mv-momentum changes abruptly, its p-momentum does not.

$$\begin{aligned} d\mathbf{p}/dt &= -q \, d\mathbf{A}/dt \\ d/dt (\mathbf{p}+q\mathbf{A}) &= d/dt (\text{p-momentum}) = 0 \end{aligned}$$

Feynman says, when magnetic fields are present, it is the p-momentum that is associated with the momentum operator  $\mathbf{P}$  and that  $(\mathbf{P}-q\mathbf{A})/m$  is the operator for velocity.

## Cooper Pairs

The phenomenon of electric current flowing without resistance derives from a remarkable behavior of electrons flowing in crystals at very low temperatures.

In V3p21-7, Feynman says:

“It turns out that due to the interactions of the electrons with the vibrations of the atoms in the lattice, there is a small net effective *attraction* between the electrons. The result is that the electrons form together, if I may speak very qualitatively and crudely, bound pairs.”

Physicists were initially quite surprised to discover that electrons can bind together in any fashion. They are negatively charged particles that repel one another, and they are identical fermions that cannot share a common state. In isolation, these facts certainly preclude electrons binding to one another. But, in the more complex environment of solid matter, other considerations apply.

In very simplified terms, as an electron moves through a crystal (a solid with a regularly repeating pattern of atoms), its negative charge attracts atomic nuclei and repels lattice electrons, resulting in a very slight shift of net positive charge, as sketched in Figure 33-3.

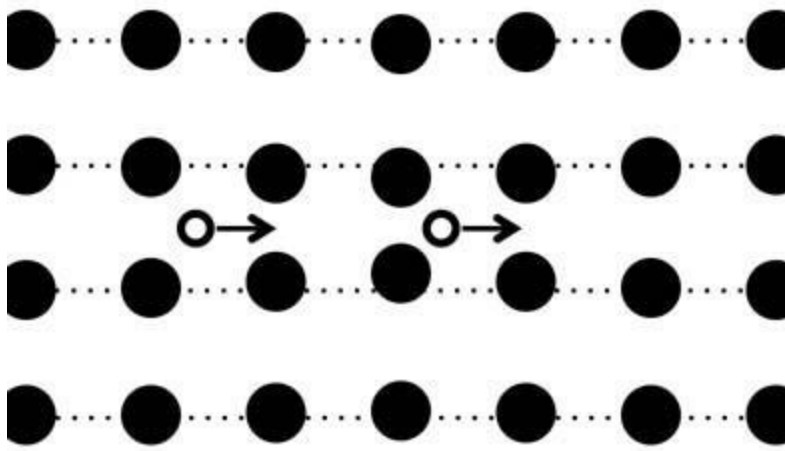


Figure 33-3 Cooper Pair Moving Through Crystal

In Figure 33-3, black dots represent net positive charges, and open circles represent electrons flowing through the crystal, with their velocities indicated by arrows. Positive charges moving toward the right-hand electron, and negative charges moving away from it (not shown in figure), create a small region of net positive charge. That positive charge attracts the left-hand electron, facilitating its passage through the crystal.

“Pairing” decreases the energy of two conduction electrons by typically 0.001 eV. This tiny energy is the effective binding energy of what is called a *Cooper pair*. Cooper pairs are unlike other bound systems. The two electrons may be thousands of atoms apart, and many Cooper pairs can overlap within the same space. Pairs continually form and dissociate. The two spin 1/2 electrons form a composite boson with a total spin of either 0 or 1.

A theoretical understanding of this remarkable phenomenon was achieved by John Bardeen, Leon Cooper, and John Schrieffer. For the *BCS theory of superconductivity*, the three authors received the 1972 Nobel Prize.

With a binding energy that corresponds to 11.6K (−261°C), Cooper pairs are easily disrupted by thermal motion, hence superconductivity exists only below a *critical temperature*,  $T_c$ , which is different for each type of superconducting material. For our purposes, we will consider a superconductor at 0K, absolute zero, to avoid the complications of dissociating Cooper pairs.

As identical bosons, Cooper pairs have an enhanced amplitude to occupy the exact same state. We found in Chapter 5 that the probability for a boson to enter a state already occupied by  $N$  other identical bosons is proportional to  $N+1$ . Hence as the temperature drops and Cooper pairs begin to accumulate in their ground state, they overwhelmingly induce all Cooper pairs to enter the exact same state.

In V3p21-8, Feynman defines  $\psi$  to be the wave function for one Cooper pair. We will factor  $\psi$  into a real function  $\rho$ , a real phase function  $\theta$ , and the normal energy exponential, as follows:

$$\psi(\mathbf{r}) = \exp\{iEt/\hbar\} \exp\{i\theta(\mathbf{r})\} \sqrt{\rho(\mathbf{r})/q}$$

The square root of charge factor makes  $\rho(\mathbf{r})$  the electric charge density, the particle density multiplied by charge.

Recall the equation for particle probability current:

$$\hat{J} = \{\psi^*(\mathbf{P}-q\mathbf{A})\psi + \psi(\mathbf{P}-q\mathbf{A})^*\psi^*\} / 2m$$

Into this equation, let's insert  $\psi$  in the above form. Since the operator  $(\mathbf{P}-q\mathbf{A})$  has no time derivatives, the factor  $\exp\{iEt/\hbar\}$  passes through the operator and cancels in  $\psi^*\psi$ . To reduce clutter, I will show the result in one dimension.

Note first that:

$$\partial/\partial x \exp\{\pm i\theta\} = \pm i\theta \partial\theta/\partial x \exp\{\pm i\theta\}$$

$$\partial/\partial x \sqrt{\rho} = (1/2\rho) \partial\rho/\partial x \sqrt{\rho}$$

$$\psi^*\psi = \rho/q$$

We now evaluate  $\hat{J}$ .

$$2m\hat{J} = \{\psi^*(-qA)\psi + \psi(-qA)^*\psi^*\} \\ + (\hbar/i) \{\psi^* \partial\psi/\partial x - \psi \partial\psi^*/\partial x\}$$

$$\{2m\hat{J} + 2\rho A\} (i/\hbar) = \\ + \psi^* (+i)\partial\theta/\partial x \psi + \psi^* \partial\rho/\partial x \psi / 2\rho \\ - \psi (-i)\partial\theta/\partial x \psi^* - \psi \partial\rho/\partial x \psi^* / 2\rho$$

$$\{2m\hat{J} + 2\rho A\} (i/\hbar) = 2(+i) \partial\theta/\partial x \rho/q$$

$$q\hat{J} = (\hbar/m) \{ \partial\theta/\partial x - (q/\hbar) A \} \rho$$

Generalizing to three dimensions, and substituting electric current  $\mathbf{J}$  for charge  $q$  times probability current  $\hat{J}$ :

$$\mathbf{J} = (\hbar/m) \{ \nabla\theta - (q/\hbar) \mathbf{A} \} \rho$$

Note that  $m$  and  $q$  are the mass and charge of a Cooper pair, each twice that of the electron.

Here,  $\nabla\theta$  is the gradient of  $\theta$ , which is the vector  $(\partial\theta/\partial x, \partial\theta/\partial y, \partial\theta/\partial z)$ . Also,  $\mathbf{J}$  is the electric current density and  $\rho$  is the electric charge density, both of which are real physical quantities that describe the superconducting electron gas. The gradient of  $\theta$  is therefore also a real measurable quantity.

By equating current density with (charge density)  $\times$  (charge velocity  $\mathbf{v}$ ), we obtain:

$$\mathbf{J} = \rho\mathbf{v} = (\hbar/m) \{ \nabla\theta - (q/\hbar) \mathbf{A} \} \rho \\ m\mathbf{v} = \hbar \nabla\theta - q \mathbf{A}$$

We see again the relationship between  $m\mathbf{v}$ -momentum and  $\mathbf{p}$ -momentum.

# Superconductivity

Ordinarily, when electric current flows through a solid, current-carrying electrons are occasionally scattered by atoms and diverted into another state. This continual loss of conduction electrons is what we call electrical resistance.

However, in superconducting materials, current is carried by Cooper pairs that are united in one common state. The amplitude for each pair to remain in that common state is extremely high, due to boson statistics. The motion of lattice atoms, muted by very low temperature, is no longer sufficient to scatter one electron out of its collective. Once initiated, an electric current will continue to flow “forever.” Superconducting currents have been observed to persist for years without measurable loss, and it is estimated that their lifetime is at least 100,000 years.

While normal metals are not superconducting above 30K ( $-243^{\circ}\text{C}$ ), exotic materials have been fabricated that are superconducting at up to 138K ( $-135^{\circ}\text{C}$ ). However, no one has yet found the Holy Grail of superconductivity: a practical material that can conduct large currents and remain superconducting at temperatures above 77K, the boiling point of liquid nitrogen. The object of this quest is fiscal rather physical. The most practical coolants are liquid nitrogen and liquid helium. When I was a graduate student, it was commonly said that liquid nitrogen cost as much as milk, while liquid helium cost as much as scotch whiskey. Both coolants now cost only one third of the corresponding beverages, but liquid helium is still 50 times the cost of liquid nitrogen. Employing superconducting on a large scale, such as AC power transmission from remote solar and wind energy farms to consumers in majors cities, remains cost-prohibitive.

## The Meissner Effect

Superconductors exclude magnetic fields from their interiors, if the fields are not too strong. If a superconductor is exposed to an external magnetic field that ramps up from zero, the changing flux density creates an electric field. That electric field drives a current in the superconductor that creates an opposing magnetic field, precisely cancelling the external magnetic field. As Feynman says: “Since all the electrons will move together, an infinitesimal electric field will generate enough current to oppose completely any applied magnetic field.”

A dramatic demonstration of magnetic field exclusion is the *Meissner effect*. Imagine a magnet resting on a superconducting metal, with everything initially at a temperature well above the superconductor’s critical temperature. If we now cool everything to below the critical temperature, a small current will flow in the superconductor, creating a magnetic field that opposes and cancels the magnet’s field within the superconductor. The opposing magnetic field levitates the magnet up off the superconductor, as shown in Figure 33-4.

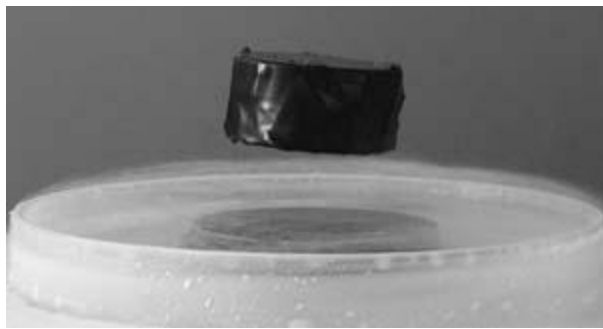


Figure 33-4 Meissner Effect  
Image by Mai-Linh Doan

Feynman shows that the Meissner effect is explained by the equations derived here.

He first adopts the standard *Coulomb gauge* in which the divergence of the vector potential  $A$  is zero. Since the electric and magnetic fields are calculated in terms of derivatives of  $A$ , the value of  $A$  itself is not fully constrained. We are free to constrain it in any manner that does not affect  $E$  or  $B$ ; most often, physicists adopt the Coulomb gauge.

Feynman says that, in any equilibrium condition, the divergence of an electric current  $J$  must be zero, “because there’s no place for it to go” — at equilibrium, current cannot continually flow into or continually flow out of any selected region. The charge density  $\rho$  must also be constant and almost completely uniform as unbalanced charges are rapidly neutralized in conducting materials. Knowing all this, we take the divergence of  $J$ .

$$\begin{aligned}\nabla \cdot J &= 0 = \{ \hbar \nabla \cdot \nabla \theta - q \operatorname{div}(A) \} \rho / m \\ 0 &= \nabla \cdot \nabla \theta = \nabla^2 \theta \\ 0 &= \nabla^2 \theta = \partial^2 \theta / \partial x^2 + \partial^2 \theta / \partial y^2 + \partial^2 \theta / \partial z^2\end{aligned}$$

In V3p21-9, Feynman says: “the only way  $\nabla^2 \theta$  can be zero everywhere inside a lump of metal is for  $\theta$  to be constant.” Mathematically,  $\nabla^2 f = 0$  does not imply  $f = \text{constant}$ ; a simple counterexample is  $f = 7x$ . Feynman’s statement is based on physics, on the nature of being “inside a lump of metal”, an electrical conductor. Unfortunately, he doesn’t explain this bit of physics here.

If  $\theta$  is constant throughout the superconductor:

$$J = -q A \rho / m$$

From the equations of electromagnetism we also have:

$$\nabla^2 A = -J / \epsilon_0 c^2$$

Combining the prior two equations yields:

$$\nabla^2 A = A \rho q / (\epsilon_0 m c^2) = A / \lambda^2$$

$$\text{with } 1 / \lambda^2 = \rho q / (\epsilon_0 m c^2)$$

The solutions for  $A$ , in one dimension are:

$$A = A_0 \exp\{-x/\lambda\}$$

Here we rejected the unrealistic choice of  $+x/\lambda$ , as that would drive the solution toward infinity. With the negative exponent, and  $x$  being the distance from the surface,  $A$  decreases exponentially from the surface toward the interior of the superconductor, with a mean attenuation length of  $\lambda$ . For a metal such as lead,  $\lambda$  is 20 nm, about 50 atomic diameters. This means only a very thin layer on the surface of the superconductor carries the current that cancels the external magnetic field.

## Flux Quantization

The London brothers, Fritz and Heinz, were two of a great many Jewish physicists (including Max Born) who fled Nazi persecution in 1933. In 1935, long before the BCS theory of superconductivity, London and London proposed their own theory of superconductivity. They predicted flux quantization: in appropriate conditions, they said magnetic flux must be an integer multiple of a specific value.

One condition involving flux quantization is illustrated in Figure 33-5. We begin with part (a) where a ring of superconducting material, warmer than its critical temperature, lies in an external magnetic field  $\mathbf{B}_{ext}$ . At this point, a magnetic field is present throughout the body of the ring.

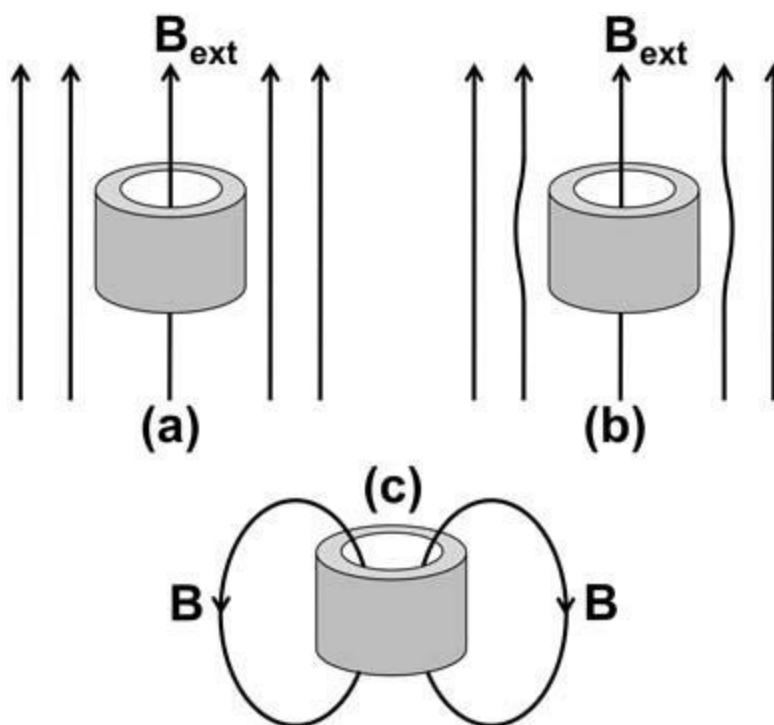


Figure 33-5 Superconducting Ring & Magnetic Field

In part (b), the ring has been cooled to well below its critical temperature. Now,  $\mathbf{B}_{ext}$  is excluded from the body of the ring by the small surface current described in the prior section, but  $\mathbf{B}_{ext}$  is still present in the hollow core inside the ring. Note that the field lines near the ring are bent outward by the cancelling magnetic field created by the superconductor.

In part (c),  $\mathbf{B}_{ext}$  has been turned off. Magnetic flux  $\Phi$  persists in the hollow core because Maxwell's equations state that  $\partial\Phi/\partial t$  is proportional to the path integral of  $\mathbf{E}$  around any circle that surrounds  $\Phi$ .

In particular, we choose path  $\Gamma$  that is completely within the superconducting material, well away from its surface, as shown in Figure 33-6. Since there can be no electric field within the interior of a superconductor, this path integral equals zero, making  $\partial\Phi/\partial t$  zero as well. As  $\mathbf{B}_{ext}$  goes to zero, a small surface current flows around the superconducting ring that maintains a magnetic field through its hollow core, permanently *trapping* the magnetic flux.

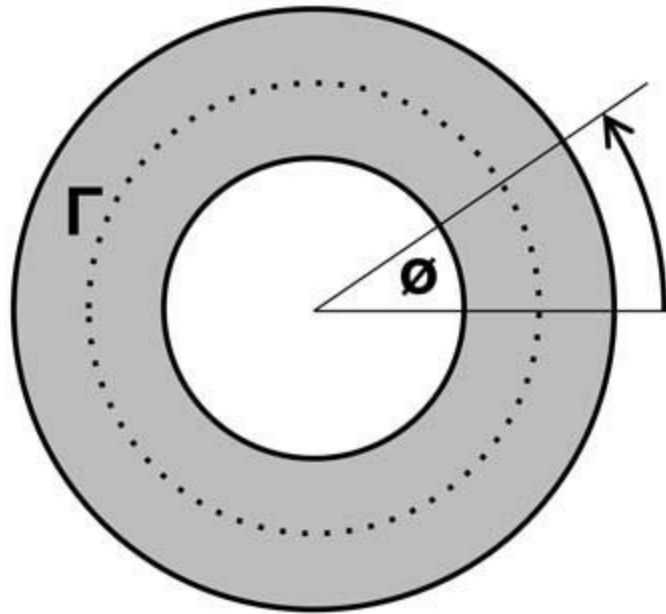


Figure 33-6 Path  $\Gamma$  Around Ring

Figure 33-6 shows a top view of the superconducting ring and path  $\Gamma$  that encircles its hollow core.

In V3p21-10, Feynman says the argument he made for  $\theta$  being constant “inside a lump of metal” works for a solid block but not for a ring. Again he provides no explanation. Let’s explore the consequences of this difference.

Recall a prior equation for the case of zero electric current:

$$0 = \mathbf{J} = (\hbar/m) \{ \check{\mathbf{D}}\theta - (q/\hbar) \mathbf{A} \} \rho$$

$$\hbar \check{\mathbf{D}}\theta = q\mathbf{A}$$

Let’s now integrate this equation along path  $\Gamma$ . We start at the 3-o’clock position in Figure 33-6, at  $\phi=0$ , and integrate counterclockwise around the circle to  $\phi=2\pi$ .

$$\hbar \int_{\Gamma} \check{\mathbf{D}}\theta \cdot d\mathbf{s} = q \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}$$

Maxwell’s equations say:

$$\int_{\Gamma} \mathbf{A} \cdot d\mathbf{s} = \Phi$$

Here  $\Phi$  is the magnetic flux passing through  $\Gamma$ . Our integral from  $\phi=0$  to  $\phi=2\pi$  becomes:

$$\int_{\Gamma} \check{\mathbf{D}}\theta \cdot d\mathbf{s} = (q/\hbar) \Phi$$

A useful theorem says that the integral of the gradient of any function  $g$  from point Y to point Z equals

$g(Z)-g(Y)$ . That theorem is more obvious if one considers its one-dimensional version:

$$\int_Y^Z \partial g(x)/\partial x \, dx = g(Z) - g(Y)$$

The result is:

$$\theta(\phi=2\pi) - \theta(\phi=0) = (q/\hbar) \Phi$$

Recall that  $\theta$  is the  $r$ -dependent phase of the wavefunction, according to:

$$\psi(\mathbf{r}) = \exp\{iEt/\hbar\} \exp\{i\theta(\mathbf{r})\} \sqrt{\rho(\mathbf{r})/q}$$

Since  $\phi=0$  and  $\phi=2\pi$  are the same point, we'd normally leap to say  $\theta(\phi=2\pi)$  must equal  $\theta(\phi=0)$ . But leap carefully; after all, this is quantum mechanics. Here, all we really know is that any wave function  $\psi$  is single-valued, meaning it has only one value at each point in space and time. Hence, the quantum requirement is:

$$\psi(\phi=2\pi) = \psi(\phi=0)$$

This will be true as long as:

$$\begin{aligned} \exp\{i\theta(\phi=2\pi)\} &= \exp\{i\theta(\phi=0)\} \\ \theta(\phi=2\pi) &= \theta(\phi=0) + 2n\pi, \text{ for any integer } n \end{aligned}$$

Combining this with the path integral of  $A$  yields:

$$\begin{aligned} 2n\pi &= (q/\hbar) \Phi \\ \Phi &= 2n\pi\hbar/q \end{aligned}$$

This means the trapped flux is quantized, always being an integer multiple of  $2\pi\hbar/q$ .

The Londons derived this formula, and using the electron's charge for  $q$ , they predicted the flux quantum was  $4 \times 10^{-7}$  gauss-cm<sup>2</sup>. For comparison, Londons' flux quantum is 1% of the flux of Earth's magnetic field through a circle with a cross-section of one human hair (0.1 mm diameter).

In V3p21-11, Feynman tells the amusing story of the search for the flux quantum:

“In 1961 such a quantized flux was looked for and found by Deaver and Fairbank at Stanford University and at about the same time by Doll and N  bauer in Germany.

“In the experiment of Deaver and Fairbank, a tiny cylinder of superconductor was made by electroplating a thin layer of tin on a one-centimeter length of No. 56 ( $1.3 \times 10^{-3}$  cm diameter) copper wire. The tin becomes superconducting below 3.8K while the copper remains a normal metal. The wire was put in a small controlled magnetic field, and the temperature reduced until the tin became superconducting. Then the external source of field was removed. You would expect this to generate a current by Lenz's law so that the flux inside would not change. The little cylinder should now have magnetic moment proportional to the flux inside. The magnetic

moment was measured by jiggling the wire up and down (like the needle on a sewing machine, but at the rate of 100 cycles per second) inside a pair of little coils at the ends of the tin cylinder. The induced voltage in the coils was then a measure of the magnetic moment.

“When the experiment was done by Deaver and Fairbank, they found that the flux was quantized, *but that the basic unit was only one-half as large as London had predicted*. Doll and Näbauer got the same result. At first this was quite mysterious, but we now understand why it should be so.”

The BCS theory explains that current in superconductors is carried by Cooper pairs, whose charge is that of **two** electrons. Therefore the true flux quantum is:

$$\Phi_0 = \pi\hbar/q_c = 2 \times 10^{-7} \text{ gauss-cm}^2$$

## Dynamics of Superconductivity

In V3p21-12, Feynman begins discussing magnetohydrodynamics, although he doesn't use that name. This is the study of the dynamics of electrically conducting fluids. Here the conducting “fluid” is a stream of Cooper pairs. Even on the Feynman scale, magnetohydrodynamics is far too esoteric for an introductory physics course. This section is a mix of partial proofs and quoted results. I will provide only the major results.

Feynman says the equations of motion for a superconducting electron fluid are:

$$\textit{curl } \mathbf{v} = -q\mathbf{B}/m$$

$$\text{with } \mathbf{v} = \{ \hbar \nabla \theta - q\mathbf{A} \} / m$$

$$m [\partial \mathbf{v} / \partial t + \mathbf{v} \cdot \nabla \mathbf{v}] = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) - \hbar^2 \nabla \{ \nabla^2 (\sqrt{\rho}) \} / 2\sqrt{\rho}$$

The term in [ ]'s in the last equation is called the fluid's *comoving* acceleration.

## Josephson Junctions

In Chapter 20, we discovered a remarkable phenomenon that occurs when different semiconducting materials are joined: the transistor. We now examine another remarkable phenomenon that occurs when joining superconducting materials: the Josephson junction. Welsh theoretical physicist B. D. Josephson predicted this effect in 1962, for which he received the 1973 Nobel Prize.

A Josephson junction consists of two superconducting regions separated by a thin insulator, as shown in Figure 33-7.

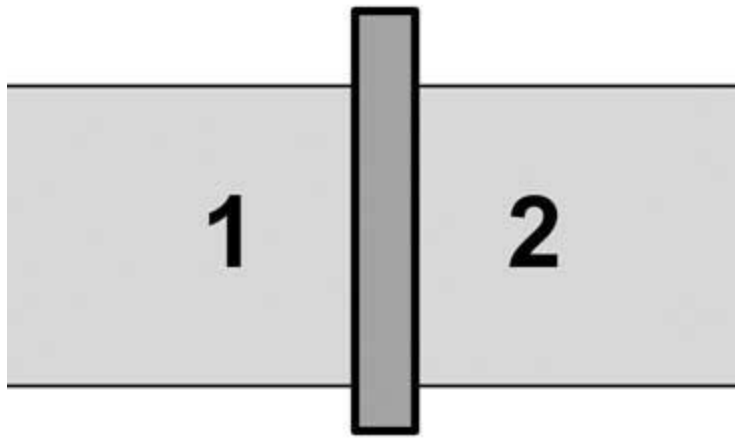


Figure 33-7 Thin Insulator  
Between Two Superconductors

If the insulator is sufficiently thin, electrons in one superconducting region can pass to the other via barrier penetration, a uniquely quantum mechanical effect.

To analyze this junction, we define  $\psi_1$  to be the wave function of Cooper pairs in region 1 and  $\psi_2$  to be the wave function in region 2. Feynman stresses that at very low temperatures, essentially all conducting electrons are in the same state  $\psi_1$  in region 1 and  $\psi_2$  in region 2.

We begin with the simplest situation. Assume both regions are made of the same material and there are no external magnetic fields. We then have a two-state system that is symmetric. The two wave functions are governed by:

$$\begin{aligned} i\hbar \partial\psi_1/\partial t &= E_1 \psi_1 + K \psi_2 \\ i\hbar \partial\psi_2/\partial t &= E_2 \psi_2 + K \psi_1 \end{aligned}$$

Here,  $K$  is the amplitude for a Cooper pair to transition from one region to the other, and  $E_1$  and  $E_2$  are the Cooper pair energy levels in each region. If the situation is completely symmetric,  $E_1=E_2$ .

Now imagine we attach the superconducting regions to opposite sides of a battery, thereby establishing:

$$E_1 - E_2 = qV$$

For convenience, we set the zero-point energy midway between  $E_1$  and  $E_2$ , such that:

$$\begin{aligned} i\hbar \partial\psi_1/\partial t &= + \psi_1 qV/2 + K \psi_2 \\ i\hbar \partial\psi_2/\partial t &= - \psi_2 qV/2 + K \psi_1 \end{aligned}$$

Next, we again factor the wave functions in terms of real functions  $\rho$  and  $\theta$ :

$$\begin{aligned} \psi_1 &= \exp\{i\theta_1\} \sqrt{\rho_1} \\ \psi_2 &= \exp\{i\theta_2\} \sqrt{\rho_2} \end{aligned}$$

In V3p21-15, Feynman provides the following results from inserting  $\psi_1$  and  $\psi_2$  into the Hamiltonian equations. I provide the derivation at the end of this chapter. To reduce clutter, define  $\delta=\theta_2-\theta_1$ . The

results are:

$$\hbar \partial \rho_1 / \partial t = +2K \sqrt{(\rho_1 \rho_2)} \sin \delta$$

$$\hbar \partial \rho_2 / \partial t = -2K \sqrt{(\rho_1 \rho_2)} \sin \delta$$

$$\hbar \partial \theta_1 / \partial t = -qV/2 + K \sqrt{(\rho_2 / \rho_1)} \cos \delta$$

$$\hbar \partial \theta_2 / \partial t = +qV/2 + K \sqrt{(\rho_1 / \rho_2)} \cos \delta$$

The two  $\partial \rho / \partial t$  equations say there is a current  $J$  flowing from region 2 to region 1, with:

$$J = J_0 \sin \delta$$

$$J_0 = (2K/\hbar) \sqrt{(\rho_1 \rho_2)}$$

This current is supplied by the external battery that maintains voltage  $V$ . We did not include that current in our calculations. A complete analysis shows that the physical characteristics of the junction ( $K$  and the  $\rho$ 's) drive the current  $J_0$ , while the battery maintains constant charge densities  $\rho_1$  and  $\rho_2$ .

The two  $\partial \theta / \partial t$  equations provide the derivative of  $\delta$ .

$$\partial(\delta) / \partial t = \partial \theta_2 / \partial t - \partial \theta_1 / \partial t = qV/\hbar$$

Integrating this equation yields:

$$\delta(t) = \delta_0 + (q/\hbar) \int V(t) dt$$

Here,  $\delta_0 = \delta(t=0)$ , and  $q$  is the Cooper pair charge. The equation for current is again:

$$J = J_0 \sin \delta$$

Feynman says the last two equations form the general theory of Josephson junctions.

For a 1 Volt potential and  $\hbar = 6.6 \times 10^{-16}$  eV-sec,  $qV/\hbar = 1/(3.3 \times 10^{-16}$  sec) and  $\sin(qVt/\hbar)$  runs through 480 trillion cycles per second. Hence the average current over any appreciable time interval is zero. Strangely, with zero applied voltage (battery replaced by a wire), a constant current flows with any value from  $-J_0$  to  $+J_0$ , as determined by the constant  $\delta_0$ . That current stops as soon as a non-zero voltage is applied, as observed by Anderson and Rowell.

Now consider applying a two-component voltage of the form:

$$V(t) = V_0 + \varepsilon \cos(\omega t), \text{ with } \varepsilon \ll V_0$$

We then have:

$$\delta(t) = \delta_0 + (q/\hbar) \{ V_0 t + \varepsilon \sin(\omega t) / \omega \}$$

Recall the equation for the sine of the sum of two angles:

$$\sin(\theta + \beta) = \sin \theta \cos \beta + \sin \beta \cos \theta$$

Now if  $\beta \ll 1$ , this reduces to:

$$\sin(\theta + \beta) = \sin\theta + \beta \cos\theta$$

We now evaluate  $\sin\{\delta(t)\}$  assuming  $q\epsilon/\hbar\omega \ll 1$ .

$$\begin{aligned} \sin\{\delta(t)\} &= \sin\{\delta_0 + (qV_0 t/\hbar)\} \\ &\quad + (q\epsilon/\hbar\omega) \sin(\omega t) \cos\{\delta_0 + (qV_0 t/\hbar)\} \end{aligned}$$

The first term on the right side of this equation averages to zero, just as described above. We can rearrange the second term to read:

$$\sin\{\delta(t)\} = (q\epsilon/2\hbar\omega) \sin\{\omega t + \delta_0 + (qV_0 t/\hbar)\}$$

The right side also rapidly averages to zero unless:

$$\omega = -qV_0/\hbar$$

This resonance phenomenon was first observed by S. Shapiro.

For the final topic of his two-year course, Feynman explores the interference between two Josephson junctions, in what we now call a SQUID, a superconducting quantum interference device. Analogous to the quantum interference in the two-slit experiment (remember *Feynman Simplified* 1A Chapter 1), when two Josephson junctions are linked in parallel, as shown in Figure 33-8, the currents through the two junctions interfere quantum mechanically.

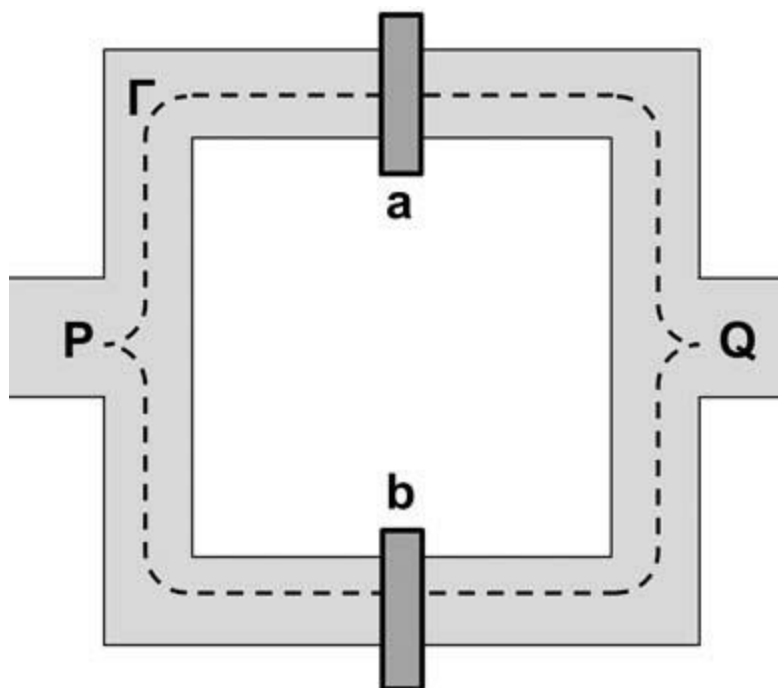


Figure 33-8 SQUID With 2 Josephson Junctions

Here  $\Gamma$  is a closed path that passes through P and Q, and encircles the hollow core of a SQUID. Detectors monitor  $J_{\text{sum}}$ , the total current flowing through the SQUID from P to Q.

Define  $J_a$  and  $J_b$  to be the current flowing through junction a and junction b. Interference arises when  $J_a$

and  $J_b$  arrive at Q with different phases.

Recall our earlier equation for the phase change along a path  $\Gamma$ :

$$\Delta\text{phase} = (q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}$$

Let  $\delta_p$  be the phase of the Cooper pairs at P. The phase at Q of the current passing through junction a, the upper half of the SQUID, is:

$$\delta_{Q_a} = \delta_p + (q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}, \text{ along upper half}$$

For the current passing through junction b, the phase at Q is:

$$\delta_{Q_b} = \delta_p + (q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}, \text{ along lower half}$$

From Maxwell's equations, we also know that the total path integral along  $\Gamma$  equals the magnetic flux  $\Phi$  through  $\Gamma$ , according to:

$$\Phi = \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}$$

Here the sign convention, per the right-hand-rule, requires traversing path  $\Gamma$  in the counterclockwise direction. The integral along the lower half of  $\Gamma$  goes counterclockwise, while the integral along the upper half of  $\Gamma$  goes clockwise. We must therefore flip the sign of the integral along the upper half of  $\Gamma$ . This yields:

$$\Phi = \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}, \text{ P to Q along lower half} \\ - \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}, \text{ P to Q along upper half}$$

$$(q/\hbar) \Phi = \delta_{Q_b} - \delta_{Q_a}$$

Here again,  $q$  is the Cooper pair charge, which equals  $2q_e$ , twice the electron charge.

For convenience, assume half the total phase shift occurs on the upper path and half on the lower path. We can then express the two currents as:

$$J_a = J_0 \sin\{\delta_p + q_e \Phi/\hbar\} \\ J_b = J_0 \sin\{\delta_p - q_e \Phi/\hbar\} \\ J_{\text{sum}} = J_a + J_b = J_0 \sin\{\delta_p\} \cos\{q_e \Phi/\hbar\}$$

While we cannot predict the value of  $\delta_p$ , we can predict that the current will be modulated sinusoidally as a function of  $\Phi$ . The current attains its greatest positive or negative value when:

$$q_e \Phi/\hbar = n\pi, \text{ for any integer } n$$

$$\Phi = n\pi\hbar/q_e$$

This is the flux quantum that we found earlier when we analyzed a superconducting ring.

SQUIDS are the most sensitive devices for measuring magnetic fields. They can detect fields as small as  $10^{-15}$  Tesla, about 10 trillionths of Earth's magnetic field.

## Josephson Wave Functions Derivation

Here is the derivation of the wave functions of Cooper pairs on opposite sides of a Josephson junction. The Hamiltonian equations are:

$$\begin{aligned} i\hbar \partial\psi_1/\partial t &= +\psi_1 qV/2 + K\psi_2 \\ i\hbar \partial\psi_2/\partial t &= -\psi_2 qV/2 + K\psi_1 \end{aligned}$$

We factor the wave functions as we did earlier for Cooper pairs:

$$\begin{aligned} \psi_1 &= \exp\{i\theta_1\} \sqrt{\rho_1} \\ \psi_2 &= \exp\{i\theta_2\} \sqrt{\rho_2} \end{aligned}$$

Recall the derivatives of these terms:

$$\begin{aligned} \partial/\partial t \exp\{i\theta\} &= i\theta \partial\theta/\partial t \exp\{i\theta\} \\ \partial/\partial t \sqrt{\rho} &= (1/2\rho) \partial\rho/\partial t \sqrt{\rho} \end{aligned}$$

$$\begin{aligned} \partial/\partial t [\sqrt{\rho} \exp\{i\theta\}] \\ = \{(1/2\rho_1)\partial\rho_1/\partial t + i\partial\theta_1/\partial t\} [\sqrt{\rho_1} \exp\{i\theta\}] \end{aligned}$$

We now put  $\psi_1$  in this form into the left side of the Hamiltonian for region 1, and separate the real and imaginary terms.

$$\begin{aligned} i\hbar \partial\psi_1/\partial t \\ &= i\hbar\{(1/2\rho_1)\partial\rho_1/\partial t + i\partial\theta_1/\partial t\} \sqrt{\rho_1}\{\cos\theta_1 + i\sin\theta_1\} \\ &= i\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \cos\theta_1 + i\partial\theta_1/\partial t \cos\theta_1 \\ &\quad + (1/2\rho_1)\partial\rho_1/\partial t(i\sin\theta_1) + i\partial\theta_1/\partial t(i\sin\theta_1)\} \\ &= i\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \cos\theta_1 - \partial\theta_1/\partial t(\sin\theta_1) \\ &\quad + (1/2\rho_1)\partial\rho_1/\partial t(i\sin\theta_1) + i\partial\theta_1/\partial t \cos\theta_1\} \\ &= i\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \cos\theta_1 - \partial\theta_1/\partial t \sin\theta_1\} \\ &\quad - \hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \sin\theta_1 + \partial\theta_1/\partial t \cos\theta_1\} \end{aligned}$$

Next, we put these expressions into the entire region 1 Hamiltonian equation, and separate the imaginary and real parts that must equate independently. This results in two equations.

$$\begin{aligned} i\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \cos\theta_1 - \partial\theta_1/\partial t \sin\theta_1\} \\ = +i\sqrt{\rho_1}\sin\theta_1 qV/2 + iK\sqrt{\rho_2}\sin\theta_2 \end{aligned}$$

$$\begin{aligned}
& -\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \sin\theta_1 + \partial\theta_1/\partial t \cos\theta_1\} \\
& = +\sqrt{\rho_1}\cos\theta_1 qV/2 + K\sqrt{\rho_2}\cos\theta_2
\end{aligned}$$

Multiply the upper equation by  $\cos\theta_1/i$  and multiply the lower equation by  $\sin\theta_1$  to get:

$$\begin{aligned}
& \hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \cos^2\theta_1 - \partial\theta_1/\partial t \cos\theta_1 \sin\theta_1\} \\
& = +\{\sqrt{\rho_1}\sin\theta_1 qV/2 + K\sqrt{\rho_2}\sin\theta_2\} \cos\theta_1 \\
& -\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \sin^2\theta_1 + \partial\theta_1/\partial t \sin\theta_1 \cos\theta_1\} \\
& = +\{\sqrt{\rho_1}\cos\theta_1 qV/2 + K\sqrt{\rho_2}\cos\theta_2\} \sin\theta_1
\end{aligned}$$

Subtracting the lower equation from the upper one yields:

$$\begin{aligned}
& \hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t\} \\
& = +K\sqrt{\rho_2}\{\sin\theta_2 \cos\theta_1 - \cos\theta_2 \sin\theta_1\} \\
& \hbar \partial\rho_1/\partial t = +2K\sqrt{(\rho_1\rho_2)} \sin(\theta_2 - \theta_1)
\end{aligned}$$

Now, going back to the real and imaginary equations from the Hamiltonian, multiply the upper equation by  $\sin\theta_1/i$  and multiply the lower equation by  $\cos\theta_1$  to get:

$$\begin{aligned}
& \hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \cos\theta_1\sin\theta_1 - \partial\theta_1/\partial t \sin^2\theta_1\} \\
& = +\{\sqrt{\rho_1}\sin\theta_1 qV/2 + K\sqrt{\rho_2}\sin\theta_2\} \sin\theta_1 \\
& -\hbar\sqrt{\rho_1}\{(1/2\rho_1)\partial\rho_1/\partial t \sin\theta_1\cos\theta_1 + \partial\theta_1/\partial t \cos^2\theta_1\} \\
& = +\{\sqrt{\rho_1}\cos\theta_1 qV/2 + K\sqrt{\rho_2}\cos\theta_2\} \cos\theta_1
\end{aligned}$$

Adding these two equations yields:

$$\begin{aligned}
& \hbar\sqrt{\rho_1}\{-\partial\theta_1/\partial t\} \\
& = +\sqrt{\rho_1} qV/2 + K\sqrt{\rho_2}\{\sin\theta_2 \sin\theta_1 + \cos\theta_2 \cos\theta_1\} \\
& \hbar \partial\theta_1/\partial t = -qV/2 + K\sqrt{(\rho_2/\rho_1)} \cos(\theta_2 - \theta_1)
\end{aligned}$$

The equations for  $\psi_2$  are simply obtained by switching the indices and the polarity of V.

$$\begin{aligned}
& \hbar \partial\rho_2/\partial t = -2K\sqrt{(\rho_1\rho_2)} \sin(\theta_2 - \theta_1) \\
& \hbar \partial\theta_2/\partial t = +qV/2 + K\sqrt{(\rho_1/\rho_2)} \cos(\theta_2 - \theta_1)
\end{aligned}$$

QED

## Finding Schrödinger's Equation For Electron In A Magnetic Field

As we did in justifying Schrödinger's equation without a magnetic field, our approach here is to

derive an equation for an electron moving along an infinite line of atoms in a magnetic field. We begin with an atomic spacing of  $b$ . Later, we will take the limit as  $b$  goes to zero and delete the atoms.

Let the atoms lie along the  $x$ -axis. For the case of no magnetic field, let the amplitude for an electron to jump from one atom to either of its immediate neighbors be  $-K$ , and let the amplitude to jump to more distant atoms be zero. As found in V2p15-9, adding a magnetic field replaces  $-K$  with:

$$-K \exp\{i(q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s}\}$$

For a line of atoms, path  $\Gamma$  is entirely along the  $x$ -axis and the path length is  $b$ . We need only concern ourselves with the  $x$ -component of  $\mathbf{A}(x)$  that we will simply call  $A$ . Define function  $f(u)$  as:

$$f(u) = (q/b\hbar) \int_{\Gamma} A dx, \text{ from } x=u-b/2 \text{ to } x=u+b/2$$

With this definition,  $u$  is the midpoint of the path. For an electron jumping from  $x-b$  to  $x$ , the integral we want is:

$$(q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s} = + b f(x-b/2)$$

For an electron jumping from  $x+b$  to  $x$ , the integration limits reverse for the path integral on the left side of the prior equation, but not on the right side;  $f(u)$  is still integrated from  $x=u-b/2$  to  $x=u+b/2$ . Here we have:

$$(q/\hbar) \int_{\Gamma} \mathbf{A} \cdot d\mathbf{s} = - b f(x+b/2)$$

We have solved the line-of-atoms problem several times before, including Chapter 18 (there the transition amplitude was  $-A$ ; here it is  $-K$ ). As before, we define  $C(x)$  to be the amplitude for an electron to be at the atom at position  $x$ . The rate of change of  $C(x)$  is given by the Hamiltonian equation:

$$\begin{aligned} i\hbar \partial C(x)/\partial t &= (E_0 + V) C(x) \\ &\quad -K \exp\{-ib f(x+b/2)\} C(x+b) \\ &\quad -K \exp\{+ib f(x-b/2)\} C(x-b) \end{aligned}$$

Here, the upper line is the energy of the electron at position  $x$ , with  $V$  being the electrostatic potential energy. The middle line is the amplitude for the electron to jump to  $x$  from  $x+b$ , and the lower line is the amplitude to jump to  $x$  from  $x-b$ .

We now take the limit as  $b$  becomes infinitesimal. When we did this before, we found that the important terms were of order  $b^2$ , so as we take this limit, we will keep terms up to order  $b^2$  and drop terms proportional to  $b^3$ ,  $b^4$ , etc.

We first take the limit of  $f(u)$ . As  $b$  goes to zero,  $\int_{\Gamma} A dx$  approaches  $A$  at the midpoint of the integral multiplied by  $\int_{\Gamma} dx=b$ . Since each exponent contains  $f$  multiplied by  $b$ , we need  $f$  up to order  $b^1$ . Let's reduce clutter a bit more by defining  $e=q/\hbar$ , and by dropping " $(x)$ " where doing so will not be confusing.

In what follows, we will repeatedly use Taylor's theorem for the series expansion of derivatives of any smooth function  $g(x)$ :

$$g(x+\Delta) = g(x) + \partial g / \partial x \Delta + \partial^2 g / \partial x^2 \Delta^2 / 2! + \dots$$

Recall the definition of  $f(u)$ :

$$f(u) = (e/b) \int_r A dx, \text{ from } x=u-b/2 \text{ to } x=u+b/2$$

As  $b$  goes to zero, the limits of  $f(x+b/2)$  and  $f(x-b/2)$  are:

$$\begin{aligned} f(x+b/2) &\rightarrow (e/b) A(x+b/2) b \\ &= e A(x+b/2) \\ &= e[A + \partial A / \partial x (b/2)] \end{aligned}$$

$$\begin{aligned} f(x-b/2) &\rightarrow (e/b) A(x-b/2) b \\ &= e A(x-b/2) \\ &= e[A - \partial A / \partial x (b/2)] \end{aligned}$$

The exponentials become:

$$\begin{aligned} \exp\{-ib f(x+b/2)\} \\ &= 1 - ibf(x+b/2) - b^2 f^2(x+b/2)/2 \\ &= 1 - ibe[A + \partial A / \partial x (b/2)] - b^2 e^2 A^2/2 \end{aligned}$$

$$\begin{aligned} \exp\{+ib f(x-b/2)\} \\ &= 1 + ibf(x-b/2) - b^2 f^2(x-b/2)/2 \\ &= 1 + ibe[A - \partial A / \partial x (b/2)] - b^2 e^2 A^2/2 \end{aligned}$$

We next approximate the  $C$ 's.

$$\begin{aligned} C(x+b) &\rightarrow C(x) + \partial C / \partial x b + \partial^2 C / \partial x^2 b^2 / 2 \\ C(x-b) &\rightarrow C(x) - \partial C / \partial x b + \partial^2 C / \partial x^2 b^2 / 2 \end{aligned}$$

Hold your breath as we put all this together. Let's work first on the middle line of the Hamiltonian equation, keeping terms up to  $b^2$ .

$$\begin{aligned} &\exp\{-ib f(x+b/2)\} C(x+b) \\ &= \{1 - ibe[A + \partial A / \partial x b/2] - b^2 e^2 A^2/2\} \\ &\quad \times \{C + \partial C / \partial x b + \partial^2 C / \partial x^2 b^2 / 2\} \\ &= C + \partial C / \partial x b + \partial^2 C / \partial x^2 b^2 / 2 - C b^2 e^2 A^2/2 \\ &\quad - (ibe)\{CA + \partial C / \partial x A b + C \partial A / \partial x b/2\} \end{aligned}$$

Now the lower line.

$$\exp\{+ib f(x-b/2)\} C(x-b)$$

$$= \{1 + i b e [A - \partial A / \partial x \, b / 2] - b^2 e^2 A^2 / 2\} \\ \times \{C - \partial C / \partial x \, b + \partial^2 C / \partial x^2 \, b^2 / 2\}$$

$$= C - \partial C / \partial x \, b + \partial^2 C / \partial x^2 \, b^2 / 2 - C b^2 e^2 A^2 / 2 \\ + (i b e) \{C A - \partial C / \partial x \, A \, b + C \, \partial A / \partial x \, b / 2\}$$

The sum of those two expressions is:

$$\exp \{-i b f(x+b/2)\} C(x+b) + \exp \{i b f(x-b/2)\} C(x-b) \\ = 2C + \partial^2 C / \partial x^2 \, b^2 - C b^2 e^2 A^2 \\ - 2(i b e) \{ \partial C / \partial x \, A \, b \} \\ = 2C + b^2 \{ \partial^2 C / \partial x^2 - C e^2 A^2 - 2(i e) \partial C / \partial x \, A \}$$

The Hamiltonian becomes:

$$i \hbar \partial C / \partial t = (E_0 + V) C - 2K C \\ - K b^2 \{ \partial^2 C / \partial x^2 - C e^2 A^2 - 2(i e) \partial C / \partial x \, A \}$$

To rewrite this, note that:

$$\{ \partial / \partial x - i e A \} \{ \partial / \partial x - i e A \} C \\ = \partial^2 C / \partial x^2 - C e^2 A^2 - 2 i e A \partial C / \partial x$$

Making this substitution, we obtain:

$$i \hbar \partial C / \partial t = (E_0 + V - 2K) C \\ - K b^2 \{ \partial / \partial x - i e A \} \{ \partial / \partial x - i e A \} C$$

Almost there!

We derived an equation for effective mass in Chapter 18, and then found in Chapter 22 that, in the limit of  $b$  going to zero and the atoms being removed, the effective mass becomes the electron's true mass. That equation, with  $K$  as the transition amplitude, is:

$$m_{\text{eff}} = \hbar^2 / (2Kb^2) \rightarrow m \\ Kb^2 \rightarrow \hbar^2 / 2m$$

Substituting this,  $e = q/\hbar$ ,  $\mathbf{P}_x = (\hbar/i) \partial / \partial x$ , and choosing  $E_0 = 2K$ , yields:

$$i \hbar \partial C / \partial t = V C \\ - (\hbar^2 / 2m) (i/\hbar)^2 (\mathbf{P}_x - qA) \{ \mathbf{P}_x - qA \} C$$

Since this equation is valid for every  $C$ , it must also be valid for the wave function  $\psi$  that is a linear combination of the  $C$ 's. Generalizing to three dimensions, we finally obtain Schrödinger's equation for an electron in a magnetic field, matching what Feynman claimed.

$$i\hbar \partial\psi/\partial t = (\mathbf{P}-q\mathbf{A}) \cdot (\mathbf{P}-q\mathbf{A}) \psi / 2m + V\psi$$

QED

## Chapter 33 Review: Key Ideas

- Superconductivity is the property of conducting electric current with *absolutely zero* resistance. From actual observations, the estimated lifetime of superconducting currents is at least 100,000 years. Theoretical estimates exceed the age of the universe.
- Many materials, but not all, are superconducting at sufficiently low temperatures.

Atomic quantum effects produce the macroscopic phenomenon of superconductivity at extreme low temperatures. This is because higher-energy states “freeze out”, and behavior is dominated by the discreteness of the remaining states. Here, Feynman says: “the quantum mechanical character of that ground state” is apparent even macroscopically.

- In the presence of magnetic fields, two types of momentum are important, each with a different role. Feynman defines these as:

“mv-momentum”:  $\mathbf{p} = m\mathbf{v}$

“p-momentum”:  $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$

Here  $\mathbf{A}$  is the vector potential defined by:

$$\mathbf{B} = \text{curl } \mathbf{A}, \text{ divergence } \{\mathbf{A}\} = 0$$

He adds that the same two types of momentum arise in classical physics as well

- In simplified terms, the *BCS theory* of superconductivity says as an electron moves through a crystal, its negative charge attracts nuclei and repels lattice electrons, resulting in a small region of net positive charge. That positive charge attracts other electrons, facilitating their passage through the crystal. Pairs of electrons are in a special sense attracted to one another in *Cooper pairs*, bound together by typically 0.001 eV. The two spin 1/2 electrons form a composite boson with a total spin of either 0 or 1. It is the Cooper pairs that carry electric current in superconductors. “Paired” electrons may be thousands of atoms apart, and many Cooper pairs can overlap within the same space. Pairs continually form and dissociate.
- The wave function for Cooper pairs and the equations for pair current and momentum are:

$$(\sqrt{\rho}) \psi(\mathbf{r}) = \exp\{iEt/\hbar\} \exp\{i\theta(\mathbf{r})\} \sqrt{\rho(\mathbf{r})}$$

$$\mathbf{J} = \rho\mathbf{v} = (\hbar/m) \{ \nabla\theta - (q/\hbar) \mathbf{A} \} \rho$$

$$m\mathbf{v} = \hbar \nabla\theta - q \mathbf{A}$$

- The flux quantum associated with superconductors is:

$$\Phi_0 = \pi\hbar/q_e = 2 \times 10^{-7} \text{ gauss-cm}^2$$

- Feynman says the equations forming the theory of Josephson junctions are:

$$\delta(t) = \delta_0 + (2q_e/\hbar) \int V(t) dt$$

$$J = J_0 \sin\delta$$

$$J_0 = (2K/\hbar) \sqrt{(\rho_1 \rho_2)}$$

- SQUIDs, superconducting quantum interference devices, are the most sensitive instruments for measuring magnetic fields. They can detect fields as small as  $10^{-15}$  Tesla, about 10 trillionths of Earth's magnetic field. The current  $J$  through a SQUID is proportional to:

$$J \sim \cos\{q_e \Phi/\hbar\}$$

- Feynman ends his course on quantum mechanics with this advice to his all-male class (unfortunately, Caltech was not coed at that time):

“I am sorry to say, gentlemen, that to participate in this adventure [physical science] it is absolutely imperative that you learn quantum mechanics as soon as possible. It was our hope that in this course we would find a way to make comprehensible to you at the earliest possible moment the mysteries of this part of physics.”

# Chapter 34

## Collapse, Entanglement & Teleportation

---

The next two chapters explore developments in quantum mechanics in the 50 years since Feynman gave his wonderful *Lectures*.

### Wave Function Collapse

The greatest unresolved mystery of quantum mechanics is *wave function collapse*.

At the 1927 Solvay conference, a triennial conclave of physics luminaries, Einstein expressed concern about the “entirely peculiar mechanism of *action at a distance*” in which quantum mechanics says wave functions instantly collapse everywhere when interactions occur.

While seemingly bizarre, wave function collapse is essential to connecting quantum mechanics with the established reality of classical scientific observation. For centuries, scientists have made countless measurements of many physical quantities, including position, mass, velocity, and temperature. Each measurement resulted in a single definite value. We have never found, for example, a football or an electron that was simultaneously in two different locations.

Quantum mechanics, however, says an electron can be in different locations simultaneously. Its wave function provides amplitudes at each different location, with corresponding probabilities that sum to 100%. But, when we measure the electron’s position, we never find it 25% here and 75% there; we always find 100% of it in exactly one location. To be consistent with the reality of measurements yielding single definite values, quantum wave functions must collapse at the instant of any substantial disturbance. In particular, at the instant of measurement, quantum mechanics must yield a single allowed value of that measurement: the wave function must instantly collapse to an eigenstate of the quantity being measured.

What triggers the collapse of a wave function is substantial interference, whether or not that interference is human-directed.

The two-slit experiment, which we discussed at the very beginning of this course, provides a vivid example of wave function collapse. As sketched in Figure 34-1, this experiment consists of a particle source S, a barrier with two slits, and a detector F. Coherent particle-waves passing through both slits combine at the detector to produce interference fringes, alternating stripes of high intensity and zero intensity.

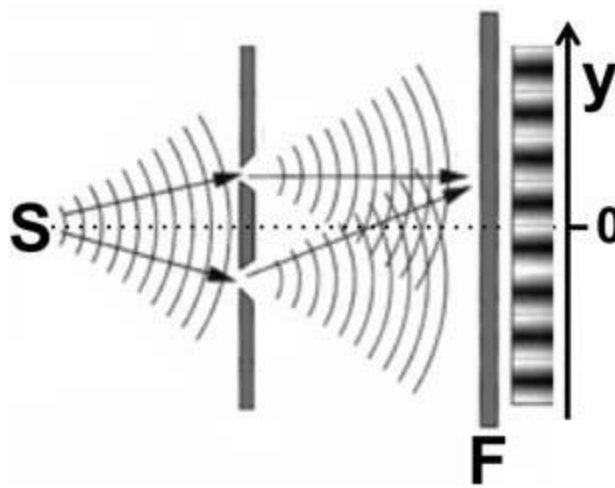


Figure 34-1 Two-Slit Experiment

The two-slit experiment is easily performed if source S emits light, but is far more difficult if S emits electrons. Dr. Tonomura accomplished the latter using an electron emission rate that allowed only one electron at a time to traverse the apparatus. In Figure 34-2, the right-hand image, with 140,000 detected electrons, clearly shows interference fringes; this proves electrons are waves. In the left-hand image, with only 200 detected electrons, the tiny dots prove electrons are particles since each strikes only a single point on the detector.

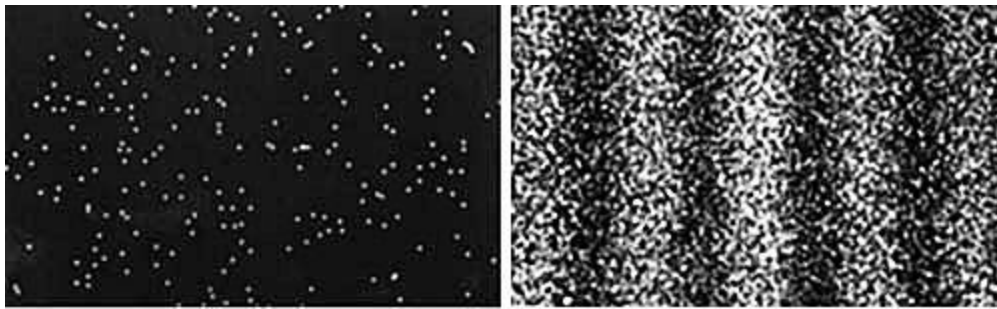


Figure 34-2 Images by Dr. Tonomura of Hitachi

The same electrons produced both images, the only difference is that there are many more of them in the right image. Since only one electron was in the apparatus at a time, the undeniable conclusion is that electrons have both particle and wave properties. Each electron must pass through both slits simultaneously and interfere with itself at F, producing a single tiny dot. But the collection of those dots is determined by the interference fringe pattern.

Figure 34-2 proves that each electron's wave function extends broadly enough to span both slits, but collapses to a tiny point as it interacts with the atoms in the detector.

At the detector, each electron's wave is at least 10,000 times wider than the tiny dots. Somehow, of the  $10^8$  spots where the electron might interact, a single impact point emerges. At that point, the interaction probability instantaneously changes from  $10^{-8}$  to 1, while the interaction probability everywhere else simultaneously becomes zero. Each electron interacts once and only once, but exactly where it will interact is completely undetermined until the instant the interaction occurs.

If wave function collapse was not universal and instantaneous, we would occasionally observe two electron strikes from a single electron wave. One might imagine that two hits could occur if they were nearly simultaneous and sufficiently separated: if a signal from the first, traveling at light speed, couldn't reach the second before its interaction. Two hits from one wave would violate the

conservation of electric charge, while communication at greater than light speed would violate special relativity. Both of those hallowed principles are even more extensively and precisely confirmed than is quantum mechanics. Indeed, no one has ever observed a violation of charge conservation or faster than light travel, even though the latter is seemingly required by wave function collapse.

To be consistent with observed reality, universal, instantaneous wave function collapse is essential, but quantum mechanics offers no explanation for the mechanism that achieves this.

After nearly a century, Einstein's quandary remains unresolved: exactly how does the wave function collapse everywhere simultaneously?

## **Interpretation: Copenhagen & Beyond**

The formalism of quantum mechanics is solid; its equations are clear, complete, and comprehensive. Its predictions have been validated by tens of thousands of high-precision experiments. The equations of quantum mechanics match what we observe in nature. The mathematical, mechanistic side of quantum mechanics is one of science's greatest triumphs. However, physicists have substantially divergent views on its interpretation — what does quantum mechanics really tell us about the fundamental processes of nature?

Newton and Einstein espoused a clear philosophy of nature. The success of their theories was established by experimental confirmation, but physicists clearly understood the world-views upon which their theories were based. In fact, Newton and Einstein established their world-views first and then derived the formalisms of their theories. Einstein, for example, postulated the constancy of the speed of light and from that derived the equations of special relativity.

Quantum mechanics is different: its formalism is clear but its world-view is fuzzy. With its indeterminacy and wave function collapse, quantum mechanics remains incompletely digested more than 100 years after its inception. We know how to do the math, but we are still searching for a clear understanding of what quantum mechanics says about nature.

So far in this course, we have presented only the standard, conventional, most widely accepted, world-view of quantum mechanics: the *Copenhagen interpretation* championed by Niels Bohr. It is the default theory to which physicists return when other ideas fail. It has achieved great success, but many physicists believe it lacks critical explanations, motivating them to intensively search for alternative interpretations.

The sheer number of alternative interpretations attests to physicists' great unease with the world-view of quantum mechanics. The most prominent interpretations, with representative dates, are: Copenhagen [1924]; pilot waves [1952]; many-worlds [1957]; modal [1969]; consistent histories [1984]; many-minds [1988]; quantum event [1989]; and relational quantum mechanics [1996]. None of the alternative interpretations has been widely accepted, at least not yet.

Through the course of this chapter and the next, we will describe the three most interesting

alternatives, starting with the most expansive.

## Many Worlds / Parallel Universes

American physicists Hugh Everett in 1957, and Bryce Seligman DeWitt in the 1960s, proposed a hypothesis that is variously called the “many worlds” or “parallel universes” theory. (Be aware that there are many other “parallel universes” theories that have nothing to do with quantum uncertainty.)

The Everett-DeWitt hypothesis is an alternative approach to the uncertainty that is central to quantum mechanics, and is perhaps best explained by considering the fate of Schrödinger’s Cat.

In this politically incorrect thought experiment, a healthy cat is sealed in a box with a radioactive source, a bottle of poison gas, and a detector that opens the bottle when the source decays. Quantum mechanics explains that radioactive sources decay at completely unpredictable moments but at precisely predictable rates. As time passes, the probability that the source has already decayed steadily increases. When that probability reaches 50%, we ask: “Is the cat alive or dead or *both*?” Conventional quantum mechanics says *both*: the cat is in a quantum superposition of two basis states,  $|\text{alive}\rangle$  and  $|\text{dead}\rangle$ , until we open the box and observe it. At that instant, the cat’s wave function “collapses” to either: (1)  $|\text{alive}\rangle$  with probability 100%; or (2)  $|\text{dead}\rangle$  with probability 100%.

Everett and DeWitt suggest this alternative: when the box is opened, our universe splits into two completely separated universes that are identical, with one exception. In one universe Schrödinger’s cat is alive and in the other it is dead. One might say the cat is both alive and dead, but in different universes, or if you prefer, in different realities.

Their approach avoids having to explain how wave functions “collapse”, but at the cost of an exponentially increasing number of parallel universes. Every time a roulette wheel stops, our universe would split into 38 new universes. And the measurement of a continuous quantity, such as momentum, would split our universe into an infinite number of new universes. Since the universe would split in response to every substantial interaction of every particle in the cosmos, the number of new universes required per second is beyond imagination. Yet, Everett and DeWitt provided no mechanism for creating this vast array of new universes.

Most physicists consider the parallel-universes / many-worlds hypothesis uneconomical, universe-wise.

## Entanglement

Entanglement is a startling consequence of quantum mechanics that physicists firmly accept, but struggle to fully comprehend.

Particles are *entangled* if they exist in a common and inseparable quantum state. We have discussed many examples of entangled states throughout this course. Two-state systems naturally produce entangled states. In Chapter 27 we discussed positronium, an “atom” composed of an electron and a

positron, its antiparticle. The energy levels of positronium are split by the interaction of the spins of the two particles. We found that the wave function of the ground state is a singlet state, an antisymmetric superposition, described by:

$$|j=0, m=0\rangle = (|e^+, p^-\rangle - |e^-, p^+\rangle)/\sqrt{2}$$

Here,  $|e^+, p^-\rangle$  is the basis state of the electron having spin up and the positron having spin down, while  $|e^-, p^+\rangle$  is the basis state of opposite spins.

The electron and positron are entangled. The two particles share one common wave function; neither has its own separate wave function. Each particle has a 50% probability to be spin up, and 50% to be spin down. But the probability of their spins being opposite is always 100%. If we measure the electron and find that it is spin up, the positron will instantaneously become spin down with 100% probability.

In V3p16-11, Feynman discusses a pair of particles 1 and 2 in an entangled state defined by  $x_1$  and  $x_2$ . He says:

“Notice that describing the position of *only one particle does not define* a [basis] state. Each [basis] state must define the condition of the entire system. You must not think that each particle moves independently as a wave in three dimensions. Any physical state  $|\psi\rangle \dots$  is therefore a function of the *two* sets of coordinates  $x_1$  and  $x_2$ . You see that such a function is not a wave in the sense of an oscillation that moves along in three dimensions. Neither is it generally simply a product of two individual waves, one for each particle. It is, in general, some kind of a wave in the six dimensions defined by  $x_1$  and  $x_2$ . If there are two particles in nature which are interacting, there is no way of describing what happens to one of the particles by trying to write down a wave function for it alone.”

The key point is: we cannot properly describe the interaction of two particles by describing only one.

As we discussed in Chapter 27, when singlet-state positronium decays, the electron and positron annihilate one another, producing two photons that move away with equal energies and in opposite directions in the CM frame. We found that the photons are always in a singlet state, an antisymmetric superposition of opposite polarizations, described by:

$$\psi/\sqrt{2} = |A=\text{RHC}, B=\text{LHC}\rangle - |A=\text{LHC}, B=\text{RHC}\rangle$$

Here,  $|A=\text{RHC}, B=\text{LHC}\rangle$  denotes photon A being RHC and photon B being LHC (left hand circularly polarized).

Now consider what happens when two scientists, Alice and Bob, equipped with photon detectors, stand on opposite sides of a stationary singlet-state positronium source. Assume that Alice is somewhat closer to the source, so that photon A reaches Alice's detector before photon B reaches Bob's detector.

The salient points seem to be:

1. Alice detects equal numbers of RHC and LHC photons, as does Bob.
2. Neither scientist can predict the polarization of the next photon: the odds are 50-50.
3. After each positronium decay, Alice and Bob always detect opposite polarizations — whenever Alice detects a RHC photon, Bob will always detect a LHC photon.
4. Surely Alice's measurement cannot affect the state of Bob's photon.

Point #4 seems to contradict point #3; hence, some say this indicates a *paradox*.

Feynman says a “paradox” isn't simply: “a conflict between reality and your feeling of what reality ‘ought to be’.” He says a “paradox” exists only when a theory makes inconsistent predictions.

For example, in *Feynman Simplified 1C* Chapter 28, we discuss the Twin Paradox of special relativity. Al remains on Earth while Bert rockets to the Delta Quadrant and back. Once reunited, each twin claims the other was moving, aged slower according to time dilation, and is now younger. Since both can't be younger, there is an *apparent* paradox. But, as we discussed, Bert had to accelerate to make a round trip, and therefore cannot claim to have been in an inertial reference frame, as required to invoke special relativity. Hence, there is no paradox; Bert is definitively younger than Al.

In fact, while there are many fascinating surprises, there are no true paradoxes in special relativity, general relativity, or quantum mechanics.

Getting back to Alice and Bob, Feynman says experiments confirm point #3: whenever Alice detects RHC, Bob is 100% certain to detect LHC. He also says point #4 is wrong: Alice's measurements *do* alter the state of Bob's photons. He says:

“[Nature] requires a description in terms of interfering amplitudes, one amplitude for each alternative. A measurement ... destroys the interference, but if a measurement is *not* made you cannot still say that ‘one alternative or the other is still occurring’.”

To expand on Feynman's statement, before the first measurement, the wave function was:

$$\psi\sqrt{2} = |A=RHC, B=LHC\rangle - |A=LHC, B=RHC\rangle$$

Here, neither photon is in a state of definite polarization. When a polarization measurement is made, quantum mechanics requires the wave function to collapse to an eigenstate of that measurement. Alice must measure either RHC or LHC, but never (RHC+LHC)/2. If she measures RHC, the wave function must have collapsed to:

$$\Phi = |A=RHC, B=LHC\rangle$$

Feynman would probably emphasize again that wave functions are mathematical constructs that have

no physical reality. When a wave function collapses throughout all of space simultaneously, no thing has actually moved; all that has changed are the probabilities of various outcomes that were previously indeterminant. While all this is entirely logical, a little voice in my head still asks how exactly does Bob's photon "know" it must now become 100% LHC.

Feynman is correct that Alice's measurements do affect the state of Bob's photons, but this wasn't properly confirmed experimentally until decades after his *Lectures*. To prove this, we must prove that the photons continue to have the original entangled wave function  $\psi$  immediately before Alice's measurement. We must exclude a possible alternative espoused by classical physics. The classical explanation is that, as the photons separate, their state is properly described by classical physics: (1) 50% of the time, photon A=RHC and B=LHC; and (2) 50% of the time, A=LHC and B=RHC. Italian physicist Carlo Rovelli says this is similar to taking two gloves and randomly sending one to Alice and one to Bob: of course the scientists always find opposite handed gloves — that's not magic, that's simply the way gloves are made.

Much more sophisticated theories and much more sophisticated experiments were required to definitively settle this "paradox", as we will discuss in the next chapter.

## Teleportation

Teleportation is the process of recreating material objects at a remote location without physically moving them through space. Teleportation is possible using entangled quantum states. While "Beam me up, Scottie!" is still just a delightful fantasy, physicists have succeeded in precisely transferring complete quantum states without physically moving them. Quantum teleportation does however require transmitting information and physically moving secondary physical objects by conventional means. Furthermore, it cannot achieve faster than light travel.

Teleportation is not a copying process that duplicates an original master. Rather, teleportation destroys the original master and creates a new master at a remote destination. If the master was information, that information no longer exists at the origin. If the original master was a material object, what remains at its former position will be an equal mass of deconstructed, disorganized, subatomic particles. The new master will have exactly the same quantum state that the original once had.

Even the most ardent Trekkie should decline an invitation to be the first teleported human or Klingon.

## Teleportation Procedure

The procedure detailed below teleports 1 bit of quantum information (1 *qubit*) from a master particle to a remote particle.

Let me explain what a *qubit* is. Conventional computers operate on *bits*; each bit can be in one of two states: 0 or 1. Digital electronic circuit designs ensure that 0's never unintentionally morph into 1's or vice versa. Integers are represented by series of bits. For example, the integer 137 can be encoded in

8 bits as 10001001 in the *binary number system*. That code represents  $2^7+2^3+2^0 = 128+8+1 = 137$ .

Quantum computers operate on qubits. Each qubit is a coherent quantum superposition of state 0 and state 1, written  $\alpha|0\rangle + \beta|1\rangle$ , where  $\alpha$  and  $\beta$  are complex numbers and  $|\alpha|^2 + |\beta|^2 = 1$ . A qubit therefore can represent any selected point on the 2-D surface of a *Bloch sphere* of radius 1, named after American physicist Felix Bloch.

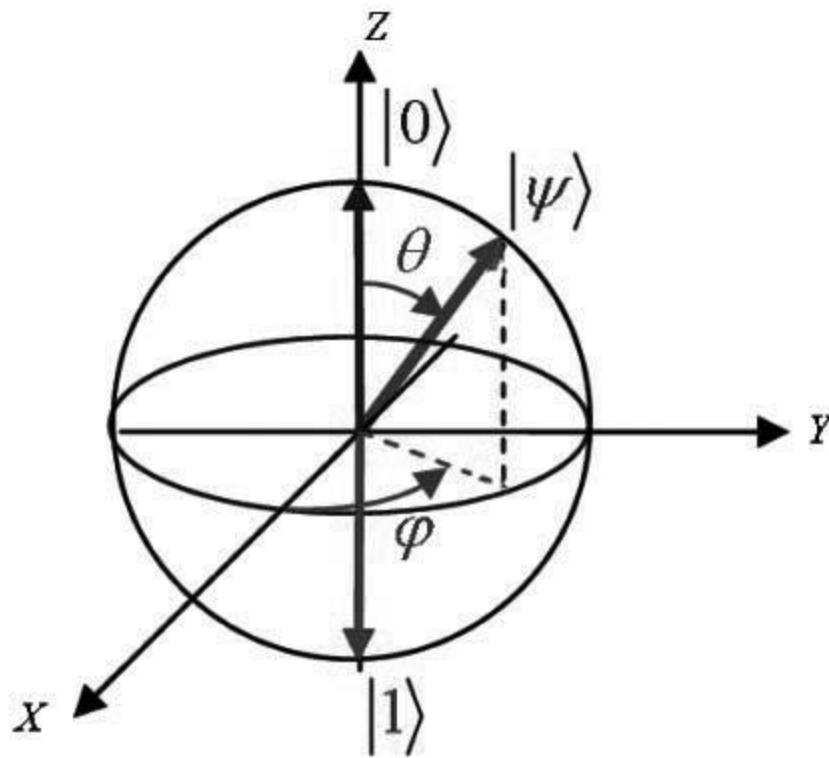


Figure 34-3 Qubit  $|\psi\rangle$  on Bloch Sphere

Figure 34-3 shows the 2-D spherical surface that represents all possible values of one qubit. The two possible values of one conventional bit are the sphere's north and south poles. The *modify* steps described below correspond to rotating the qubit across the surface of the Bloch sphere without disrupting its coherence.

Eight conventional bits, called one *byte*, can represent **any** number from 0 to 255. Eight qubits, a *qubyte*, can simultaneously represent **every** number from 0 to 255, for as long as quantum coherence persists.

Performing a computational procedure on one qubyte effectively performs that procedure simultaneously on all superposition combinations, vastly increasing computing power. This is the goal of quantum computing, which is now beginning to be realized.

In the simplest example of teleportation, the master originally contains 1 qubit of information that is a superposition of two complementary states. For specificity, we'll assume the complementary states are spin +S and -S, although the procedure is also applicable to other situations.

The truly clever aspects of this procedure are the *modify* steps devised by Canadian physicists Charles Bennett, et. al., and published in 1993.

## 6-Step Teleportation Procedure

1. Make entangled particle pair BC.
2. Send C to destination by normal means.
3. Conditionally *modify* B per state of master.
4. Measure z-spin of B and x-spin of master.
5. Transmit those spin values to destination.
6. Conditionally *modify* C; C is now master.

In Step 1, particles B and C are an entangled particle pair in a superposition of two complementary spin states. We will assume that master, B, and C are identical particles — three electrons, for example. Teleportation can be accomplished if B and C are identical particles and master is a different particle, but that requires additional steps.

In Step 2, C is physically transported to the destination by conventional means. According to special relativity, this cannot occur at more than light speed. To reduce communication delays, this transfer could occur well before the decision to teleport information, but B and C must remain isolated from external disturbances to maintain their entanglement.

In Step 3, the z-spin component of B is flipped conditionally. This operation, called *CNOT* for *conditional negation*, flips B's spin (exchanging up and down) if the z-spin component of master is down, and does nothing if the z-spin component of master is up. The trick is performing CNOT without measuring master's spin, which would destroy essential quantum coherence. Perhaps surprisingly, certain quantum devices can perform *unitary operations* such as CNOT without a direct "measurement" and without loss of coherence — this is the real "magic." A unitary operation is equivalent to rotating the qubit to another point on the Bloch sphere without "touching" it. Performing CNOT on B does not change the state of C, because CNOT does not collapse any wave functions.

In Step 4, we intrusively measure two spin components: B's z-spin and master's x-spin. This results in four possible outcomes (B up/down  $\times$  master left/right), which amounts to two conventional bits. The states of both B and master are completely disrupted by the conventional measurement process. The particle that was master still exists, but the essential characteristics that made it worth describing are irrevocably lost.

In Step 5, the two measured spin values are transmitted to the destination, at no more than light speed.

In Step 6, the state of C is conditionally *modified* according to the spin measurements made on B and master. The possible *modifications* are 180° rotations on the Bloch sphere about either the x- or z-axis. Rotating about x flips up and down, and rotating about z flips left and right. These rotations, also unitary operations, must be performed without measuring particle spins, which would destroy quantum coherence. The rotations, if required, must be performed in this order:

- a) If master's spin was left, flip C left/right.
- b) If B's spin was up, flip C up/down.

All this results in master's original quantum state being transferred to C at the remote destination.

## Proof of Teleportation Procedure

This section provides mathematical proof of the procedure's validity using the following notation for quantum states:

$|U\rangle$  means spin up (+z)  
 $|D\rangle$  means spin down (-z)  
 $|R\rangle$  means spin right (+x)  
 $|L\rangle$  means spin left (-x).

For simplicity, I won't show the overall normalization of each wave function, which is determined by requiring  $|\psi| = 1$ .

First, we need to review some spin basics. From the spin rotation tables in Chapter 17, the relationship between z-spins (up/down) and x-spins (right/left) is:

$$\begin{aligned}|U\rangle &= |R\rangle + |L\rangle \\ |D\rangle &= |R\rangle - |L\rangle\end{aligned}$$

The result of rotating z-spins by  $180^\circ$  about z-axis is:

Start, any U/D:  $a|U\rangle + b|D\rangle$   
Restate in R/L:  $a|R\rangle + a|L\rangle + b|R\rangle - b|L\rangle$   
Rotate about z:  $a|L\rangle + a|R\rangle + b|L\rangle - b|R\rangle$   
Restate in U/D:  $a|U\rangle - b|D\rangle$

Net effect: invert polarity of all  $|D\rangle$  states.

We are now ready to confirm the teleportation procedure. We will deal with quantum states of three identical particles. Let  $|L,U,D\rangle$  denote master has spin left, B has spin up, and C has spin down. Our objective is teleporting the initial state of master to the remote particle C.

Initial states after Step 1:

$$\begin{aligned}\text{BC: } \psi &= |U,D\rangle - |D,U\rangle \\ \text{Master: } \psi &= \alpha|U\rangle + \beta|D\rangle\end{aligned}$$

Initial 3-particle combined state:

$$\psi = +\alpha|U,U,D\rangle - \alpha|U,D,U\rangle + \beta|D,U,D\rangle - \beta|D,D,U\rangle$$

**Step 3:** We apply CNOT: flipping B if master is down. This yields:

$$\psi = +\alpha|U,U,D\rangle - \alpha|U,D,U\rangle + \beta|D,D,D\rangle - \beta|D,U,U\rangle$$

As Step 4 measures master's x-spin, we restate  $\psi$  in terms of master spin states left and right:

$$\begin{aligned}\psi = & \\ & +\alpha|R,U,D\rangle -\alpha|R,D,U\rangle +\beta|R,D,D\rangle -\beta|R,U,U\rangle \\ & +\alpha|L,U,D\rangle -\alpha|L,D,U\rangle -\beta|L,D,D\rangle +\beta|L,U,U\rangle\end{aligned}$$

Step 4: we measure the x-spin of master and z-spin of B, with 4 possible outcomes. Let S=L,U denote master's spin was left and B's spin was up, etc. Note that entanglement is built into  $\psi$ ; a measurement finding B is up eliminates all states with B down, but does not change the spin of C in any remaining states. The four possible outcomes of Step 4 are:

$$\begin{aligned}S=R,U: \psi &= +\alpha|R,U,D\rangle -\beta|R,U,U\rangle \\ S=L,U: \psi &= +\alpha|L,U,D\rangle +\beta|L,U,U\rangle \\ S=R,D: \psi &= -\alpha|R,D,U\rangle +\beta|R,D,D\rangle \\ S=L,D: \psi &= -\alpha|L,D,U\rangle -\beta|L,D,D\rangle\end{aligned}$$

Step 5: we transmit S to the destination by conventional means.

Step 6a: flip C right/left if master was right. As we saw above, this operation on C up/down states flips the sign of C down states. In each of the four possible S outcomes, this yields:

$$\begin{aligned}S=R,U: \psi &= -\alpha|R,U,D\rangle -\beta|R,U,U\rangle \\ S=L,U: \psi &= +\alpha|L,U,D\rangle +\beta|L,U,U\rangle \\ S=R,D: \psi &= -\alpha|R,D,U\rangle -\beta|R,D,D\rangle \\ S=L,D: \psi &= -\alpha|L,D,U\rangle -\beta|L,D,D\rangle\end{aligned}$$

Step 6b: we next flip C up/down if B was up. Recall that S=L,U denotes master's original x-spin was left and B's z-spin was up, etc. In each of the four possible S outcomes, this yields:

$$\begin{aligned}S=R,U: \psi &= -\alpha|R,U,U\rangle -\beta|R,U,D\rangle \\ S=L,U: \psi &= +\alpha|L,U,U\rangle +\beta|L,U,D\rangle \\ S=R,D: \psi &= -\alpha|R,D,U\rangle -\beta|R,D,D\rangle \\ S=L,D: \psi &= -\alpha|L,D,U\rangle -\beta|L,D,D\rangle\end{aligned}$$

In each of the four possible S outcomes, the final quantum state of C is:

$$C: \psi = \pm(\alpha|U\rangle + \beta|D\rangle) = \pm(\text{initial state of master})$$

As we know, multiplying  $\psi$  by  $\exp[i\theta]$  for any angle  $\theta$  (0 or  $\pi$  in this case) changes  $\psi$ 's phase angle but not its magnitude. Phase angles are important only when wave functions interfere coherently. In isolation, phase angles are irrelevant. Therefore, the above procedure has indeed teleported a master particle's initial quantum state to a remote particle, but at the cost of disrupting the state of master, transmitting 2 bits conventionally, and physically transporting half of an entangled pair.

## Teleportation Considerations

Quantum teleportation, QT, cannot reach destinations that have not previously been reached by conventional means. QT cannot teleport space explorers to distant planets until after halves of an immense number of entangled pairs and considerable equipment have been delivered to those destinations. Even then, one entangled pair is consumed for every quantum state teleported. Transporting entangled particles while maintaining their coherence is much more difficult than transporting conventional materials. These issues will likely severely restrict the applications of QT.

Teleportation is more than a mere pipedream. QT has been accomplished to a limited degree by the above procedure. So far, a few qubits, but no physical objects, have been teleported. Error rates are declining, but are still too high for practical applications. The greatest teleportation distance to date is 144 kilometers (between the Canary Islands). That distance is comparable to the altitude of orbiting satellites, proving the possibility of worldwide teleportation of quantum information.

Teleporting a person would require transferring the quantum states of at least  $10^{28}$  particles and then reassembling those particles. Hopping a plane, or a spacecraft, seems more practical. One might also ask: “Is there anything in a person beyond the sum of their particle states?” Scientists aren’t aware of anything additional, but would you bet your life on it?

## Communicating on the QT

Quantum teleportation will enable communication that is vastly more secure than conventional alternatives. Currently, secure communications rely on senders and receivers knowing secret encryption keys for translating encoded messages. However, such keys can be copied by, leaked to, or deciphered by adversaries. During World War II, for example, the British cracked Germany’s Enigma code and the Allies captured Enigma machines. The U.S. also cracked Japan’s naval codes. Germany and Japan paid dearly for being overly confident in their communication security.

In contrast, the “encryption key” of quantum communication is the physical possession of half of an entangled particle pair, which unless meticulously isolated, decoheres destroying the key. Nature’s quantum uncertainty guarantees message scrambling is absolutely random, and therefore undecipherable.

One major disadvantage of quantum cryptography is that each entangled pair can only teleport one qubit from one sender to one receiver. It might not be adequate to communicate lengthy, urgent, or broadcast messages.

## Chapter 34 Review: Key Ideas

- **Wave function collapse** is the greatest unresolved mystery of quantum mechanics. In 1927, Einstein described it as the “entirely peculiar mechanism of *action at a distance*.” Quantum mechanics says objects can exist in a superposition of different momenta, spins, locations, etc., but measurements always yield single definitive values for these quantities. To reconcile quantum indeterminacy with definitive measurements, quantum wave functions must collapse everywhere simultaneously when

disturbed by any substantial external influence, including scientific observation. No one understands the “magical” mechanism by which wave functions actually do collapse.

- The mathematical formalism of quantum mechanics is one of science’s greatest triumphs. However, physicists have substantially divergent views on its interpretation. What is the world-view of quantum mechanics — **what does quantum mechanics really say about the fundamental processes of nature?**

This course has previously presented only the standard, most widely accepted, world-view of quantum mechanics: the *Copenhagen interpretation* championed by Niels Bohr. It has achieved great success, but many physicists find it lacking critical explanations, motivating them to intensively search for alternative interpretations. The sheer number of alternative interpretations attests to great unease with the world-view of quantum mechanics. The most prominent interpretations, with representative dates, are: Copenhagen [1924]; pilot waves [1952]; many-worlds [1957]; modal [1969]; consistent histories [1984]; many-minds [1988]; quantum event [1989]; and relational quantum mechanics [1996]. None of the alternative interpretations has been widely accepted, at least not yet.

- Particles are *entangled* if they exist in a common and inseparable quantum state. For example, the ground state of positronium is:

$$|j=0, m=0\rangle = (|e^+, p^-\rangle - |e^-, p^+\rangle) / \sqrt{2}$$

Here,  $|e^+, p^-\rangle$  is the basis state of the electron having spin up and the positron having spin down, and  $|e^-, p^+\rangle$  is the basis state of opposite spins. The two particles share one common wave function; neither has its own separate wave function. Each particle has a 50% probability to be spin up, and 50% to be spin down. But the probability of their spins being opposite is always 100%. If we measure the electron and find that it is spin up, the positron will instantaneously become spin down with 100% probability.

- Teleportation is the process of recreating material objects at a remote location without physically moving them through space. Teleportation is possible using entangled quantum states. Quantum teleportation cannot reach destinations that have not previously been reached by conventional means. Even then, one entangled particle pair is consumed for every quantum state teleported. Transporting entangled particles while maintaining their coherence is much more difficult than transporting conventional materials.

# Chapter 35

## EPR: QM vs. Local Realism

---

The story of entanglement begins with Albert Einstein, as does much of modern physics. In 1935, Einstein and two junior associates, Boris Podolsky and Nathan Rosen, published a paper intending to demonstrate the inadequacy of quantum mechanics. Known as the *EPR* paper after the authors' initials, this is one of the most famous and most important contributions to the debate, development, and understanding of quantum mechanics. The complete title is "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?"

EPR asserted that a comprehensive physical theory must address everything that is *real*. They further stated that a physical quantity must be considered *real* if it can be predicted with absolute certainty, without disturbing the system of which it is a part. For example, if a particle's position is precisely determined even without being measured, that position is a *real* physical entity and any complete theory of physics must definitively describe it. They claimed quantum mechanics, with its Uncertainty Principle, fails this essential requirement.

EPR proposed a thought experiment that no one else had seriously considered: measurements of entangled particle pairs.

Imagine two entangled particles, A and B, in a known state, moving away from point Q in opposite directions. At some point, when their separation is sufficiently large, we will assume the particles can no longer interact with one another.

If we then precisely determine the momentum of A, that act precisely and immediately determines the momentum of B, due to the conservation of total momentum. By EPR's definition, the act of measuring A's momentum makes B's momentum *real*, even without measuring B. But, according to the Uncertainty Principle, precisely knowing B's momentum makes its location completely uncertain.

Conversely, if we instead precisely determine the position of A, that act precisely and immediately determines the position of B, since their common center of mass is stationary. By EPR's definition, the act of measuring A's position makes B's position *real*, even without measuring B. But, again, precisely knowing B's position makes its momentum completely uncertain.

But, if the particles don't interact after separating, how can the reality of B's properties depend on how we choose to measure A (its momentum or its position)? To Einstein, this was irrational. This dilemma is called the *EPR paradox*.

EPR claimed there were only two possible resolutions of this paradox:

1. contrary to our assumption, particles can interact instantaneously, even when widely separated;  
or
2. particle positions and momenta are always real and definite, independent of our measurements.

Alternative (1) violates the postulate of special relativity that nothing can move through space faster than the speed of light. Alternative (2) violates the Uncertainty Principle of quantum mechanics.

Einstein rejected alternative (1), and concluded that quantum mechanics does not fully describe reality and is therefore **incomplete**. Einstein believed quantum mechanics provides valuable insights, but is only an intermediate step toward a more comprehensive understanding of nature's fundamental truths.

EPR suggested the possibility of underlying, but as yet undiscovered, *elements of reality* that precisely determine the physical quantities that quantum mechanics says are inherently indeterminant, such as position and momentum. These undiscovered *elements of reality* became known as *hidden variables*. The basic idea is: nature seems uncertain only because we don't see the whole picture — if the hidden variables were revealed, we would clearly see that nature is devoid of uncertainty and completely deterministic.

Pre-quantum physicists believed in *Locality* and in *Realism* (the combination of both is called *Local Realism*).

**Locality** denies action-at-a-distance, particularly instantaneous action-at-a-distance: particles, energy, forces, and information can flow from one location to another, but not faster than the speed of light. An effect cannot precede the arrival of a signal traveling at light speed from its cause.

**Realism** demands that physical quantities, such as a particle's energy, are real and precisely determined, independent of their interaction with anything else, including our instruments.

**Local Realism** is implicitly assumed by classical (pre-quantum) science.

What had been a collegial philosophical debate was transformed by EPR into concrete, stark physical terms, creating an existential confrontation between the two great pillars of 20th century physics: relativity and quantum mechanics.

## The EPRB Formulation

Einstein's focus on complementary measurements of entangled particle pairs was brilliant. But the particular formulation presented by EPR was problematic, particularly regarding experimental testing. David Bohm developed a better formulation of Einstein's concept. With Yakir Aharonov, Bohm published his completed version in 1957, 22 years after the EPR paper. This formulation, called EPRB, has been the basis of all subsequent work on the EPR paradox.

Bohm's EPRB formulation is what we discussed earlier: two identical entangled particles, in a singlet state with opposite spins, move far apart, and are then subjected to spin measurements. We call the particles A and B, and denote the spin states U for up and D for down. Assume again that A is measured before B.

Quantum mechanics says the entangled singlet wave function is a quantum superposition, described by:

$$\psi/\sqrt{2} = |A=U, B=D\rangle - |A=D, B=U\rangle$$

Prior to any measurement, the spin of each particle is indeterminant: each is equally U and D. However, the act of measuring A's spin instantaneously alters B's spin, immediately forcing it to become the opposite of A's, thus ensuring total spin zero.

In the Bohm formulation, the paradox becomes: (1) are both particle spins actually well-determined before the measurement of A, contrary to quantum mechanics; or (2) if quantum mechanics is correct, how does the measurement of A dictate the state of B?

## Bohm's Pilot Waves

Bohm went further and developed an innovative alternative to conventional quantum mechanics. The standard interpretation of particle-wave duality is that every entity *has*, to varying degrees, both wave properties and particle properties — everything is both a particle ***and*** a wave. In Bohm's alternative, everything is a particle ***plus*** a wave. Every entity, Bohm said, has two parts: one part that behaves like a classical particle and another part that behaves like a classical wave. (Louis de Broglie had suggested a similar notion many years earlier, but quickly abandoned it.) An electron, Bohm said, is a classical particle plus an associated *pilot wave*. The pilot wave spreads throughout space “sensing” its environment, after which it somehow appropriately directs the motion of its associated particle. In some versions of Bohm's scheme, faster than light speed is not necessary, because pilot waves sense the environment as it develops, even before physicists start collecting data.

In the basic two-slit experiment, for example, each pilot wave explores the geometry of the apparatus and directs its particle toward a spot on the detector. After myriad particles are piloted through the apparatus, interference fringes appear on the detector. Exactly where each particle hits the detector is still governed by probability; God still plays with dice. But the difference Bohm proposes is that each particle follows a predefined path from source to detector. The paths are assigned randomly, according to a probability distribution determined after its pilot wave senses the environment but before the particle leaves the source.

By analogy, consider a deck of playing cards. We shuffle the cards and then deal them. In Bohm's concept, each card ends up in a random place, but a place that was determined before the card was dealt. By examining the deck before the first card is dealt, we can discover the final destination of every card. Conversely, according to conventional quantum mechanics, each card finds its own path to a random destination. Examining the deck before it is dealt provides no information about where each card will land. In quantum mechanics, shuffling is unnecessary; randomization occurs as each

card flies through the air. It's much harder for a quantum dealer to cheat.

If an external disruption is added to the two-slit apparatus, such as short-wavelength light shining on one slit, pilot waves would re-direct their particles and the interference fringes would vanish.

In the EPRB scenario, Bohm says:

1. Particles in singlet states always have opposite spin.
2. The spin axis of each particle pair is set at the source.
3. The spin axis doesn't change from source to detector.
4. Each particle always spins in a well-defined direction.

Naturally, Bohm says, measurements on separated pairs will discover opposite spins, regardless of how and where the spins are measured.

Bohm said pilot waves have so far escaped detection because their energies are vastly smaller than the energies of their associated particles. In his book *The Undivided Universe*, Bohm compares a pilot wave to a radar beacon leading a ship toward port through thick fog — the ship might have trillions of times more energy than the beacon, but follows it nonetheless.

Could Bohm's pilot waves be Einstein's hidden variables?

## **EPR Tests: QM vs. Local Realism**

How can physicists determine whether Local Realism or quantum mechanics best describes nature?

The conflict may be clarified with this macroscopic analogy. Recall the gloves in the prior chapter. Imagine I order a pair of gloves to be delivered in two boxes, one glove in each box. Not knowing which is which, I mail one to Alice and the other to Bob. Since human intuition is grounded in Local Realism, no one is surprised that Alice and Bob receive opposite-handed gloves. But what if these were quantum gloves in a singlet state? Quantum mechanics says each glove has half a thumb on each side as long as they're boxed. When the first box is opened, one glove randomly and instantly becomes totally left-handed and the other automatically and simultaneously becomes totally right-handed. While the final result is the same in both theories, the quantum description of how it happens profoundly shocks human intuition.

In the EPRB formulation, Local Realism says each particle always has a fixed well-defined spin, and paired singlet particles have opposite spins. Conversely, quantum mechanics says the particles are entangled, neither particle has a well-defined spin before an external interaction, such as a measurement, but their total sum is always definitively zero.

In both theories, singlet particle spins are always measured to be exactly opposite one another. How can we distinguish between two theories that predict the same outcome? How can we know the true state of particle spins before measuring them?

# BELL'S INEQUALITY

In 1964, Irish physicist John Stewart Bell found the path forward. He devised *Bell's Inequality Test* to prove whether or not particle spins are well-determined before they are measured. His clever idea requires measuring how particle spins correlate when measured in different directions.

For specificity, let's assume the singlet particles are photons with spin 1. Spin-measuring devices will report their spins as either +1 or -1 along any direction we choose, but each measurement is performed only along one direction. A device measuring spins vertically, as shown on Alice's side of Figure 35-1, will report either: +1 corresponding to +z-spin, or -1 meaning -z-spin. Another measuring device rotated by angle  $\theta$  about the y axis, as shown on Bob's side, will report either: +1 meaning along the dashed arrow, or -1 meaning antiparallel to the dashed arrow.

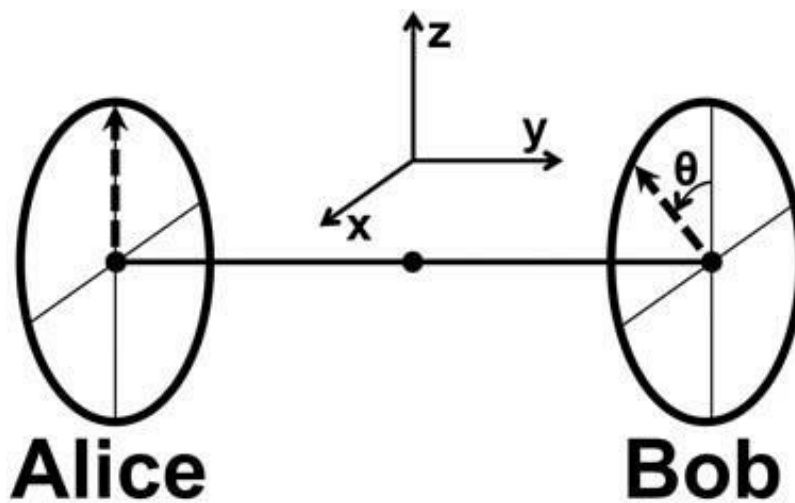


Figure 35-1 Orientation of Spin Measurements

We will keep  $\theta$  in the range  $-180^\circ$  to  $+180^\circ$ .

Quantum mechanics and classical physics both agree that spin-measuring devices oriented in the same direction ( $\theta=0$ ) will report opposite spin values: we will always find one +1 and one -1. But things get more interesting when the devices are pointing in different directions, when  $\theta$  is non-zero.

Now imagine viewing all this by looking along the horizontal y-axis, as shown in Figure 35-2. On the left, Alice's detector reports +1 if her particle's spin direction, the dotted arrow, is within the gray zone that spans all angles within  $90^\circ$  of her detector's measurement direction, the solid arrow. Conversely, her detector reports -1 if the spin is in the white zone that spans angles more than  $90^\circ$  from her detector orientation.

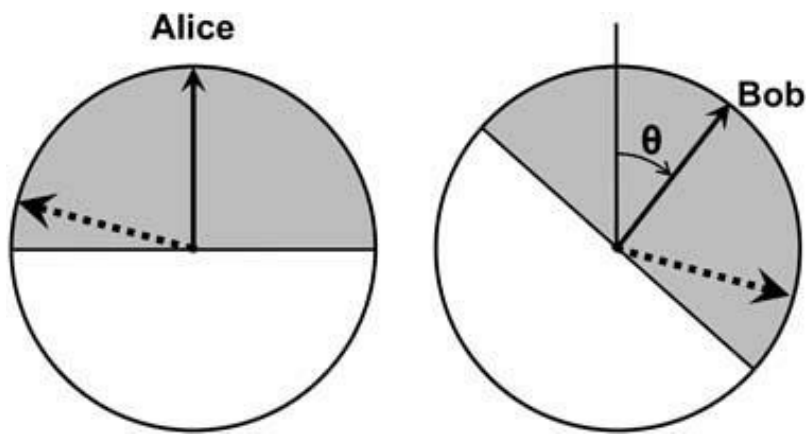


Figure 35-2 Detector Responses

On the right, Bob's detector reports +1 if his particle's spin direction, the dotted arrow, is within the gray zone that spans all angles within  $90^\circ$  of his detector's measurement direction, the solid arrow at angle  $\theta$  relative to Alice's detector orientation. His detector reports -1 if the spin is in the white zone, any direction more than  $90^\circ$  from his detector orientation.

Now consider what the entire apparatus looks like in the combined view shown in Figure 35-3.

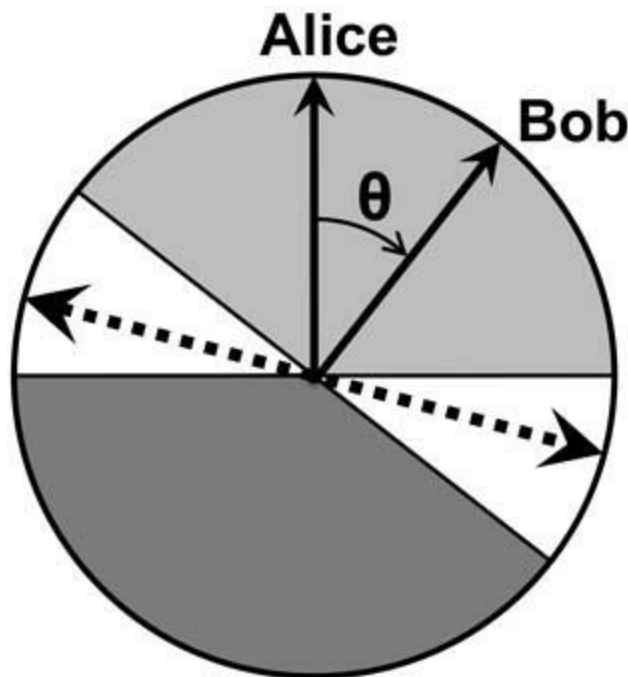


Figure 35-3 Combined Detector Response Zones

Here Alice's detector is oriented toward +z, as before, and Bob's detector is rotated clockwise by an angle  $\theta$ . Be aware that the definitions of grey and white zones in Figure 35-2 and Figure 35-3 are different.

In classical physics, with or without hidden variables, particle pairs leave the origin spinning in opposite directions along the same axis, as represented by the dotted arrows. The orientation of that axis is random. A spin in the light gray zone would be reported +1 by both Alice and Bob. A spin in the dark gray zone would be reported -1 by both. Spins in the white zones would be reported +1 by one scientist and -1 by the other. The four possible combinations, with measured spins and the product of the two measured spins are:

Alice's in light gray: +1

Bob's in dark gray:  $-1$   
Spin product:  $-1$

Alice's in dark gray:  $-1$   
Bob's in light gray:  $+1$ .  
Spin product:  $-1$

Alice's in upper white:  $+1$   
Bob's in lower white:  $+1$   
Spin product:  $+1$

Alice's in lower white:  $-1$   
Bob's in upper white:  $-1$   
Spin product:  $+1$

Each white zone spans an angle  $\theta$ , and each gray zone spans an angle  $\pi-\theta$ . Spin correlation  $C$ , defined as the average over all spin orientations of the product of the two reported spins divided by the total spin squared (1 in this case), is:

$$C(\theta) = \{-2(\pi-\theta) + 2\theta\} / 2\pi$$
$$C(\theta) = -1 + 2\theta/\pi$$

This is the prediction of classical physics, with or without hidden variables. For simple cases, it matches what we expect. For  $\theta=0$ , both devices are oriented toward  $+z$  and  $C=-1$ ; spins are found opposite 100% of the time. For  $\theta=\pi$ , the devices are oppositely oriented and  $C=+1$ ; spins are found equal 100% of the time. For  $\theta=\pi/2$ , the devices are perpendicular and  $C=0$ ; the spins are uncorrelated — regardless of Alice's spin, Bob's spin is just as likely to be  $+1$  as  $-1$ .

What happens in quantum physics, where particle spins are not definite until measured? Assume Alice's spin is measured first and is reported  $+1$ , which is  $+z$  in the above figure. Bob's spin immediately becomes  $-z$ . If  $\theta=0$ , Bob's detector reports  $-1$ .

But, what does Bob's detector report when  $\theta$  is non-zero?

For spin  $-1$  along the  $z$ -axis, the amplitudes for spin  $-1$  and  $+1$  along an axis rotated from the  $z$ -axis by angle  $\theta$  about the  $y$ -axis are (see rotation tables in Chapter 17):

$$\langle -1 \text{ at angle } \theta \mid -z \rangle = (1+\cos\theta)/2$$
$$\langle +1 \text{ at angle } \theta \mid -z \rangle = (1-\cos\theta)/2$$

The expectation value for Bob's spin measurement equals:

$$\begin{aligned} & -1 \times \{\text{Prob}(\text{spin}=-1) + 1 \times \{\text{Prob}(\text{spin}=+1) = \\ & \quad -1 \times (1+2\cos\theta+\cos^2\theta)/4 \\ & \quad +1 \times (1-2\cos\theta+\cos^2\theta)/4 \\ & = -\cos\theta \end{aligned}$$

The expectation value of the product of Alice's and Bob's measured spins is:

$$C(\theta) = (\text{Alice's spin}) \times (\text{Bob's spin})$$

$$C(\theta) = (+1) \times (-\cos\theta)$$

$$C(\theta) = -\cos\theta$$

This is the result when Alice measures spin +1. You may wish to repeat this analysis for Alice measuring -1; the result is the same.

Quantum mechanics and classical physics make the same predictions for  $\theta = 0, \pm 90^\circ$ , and  $\pm 180^\circ$ . But at other angles, their predictions differ slightly, as shown in Figure 35-4.

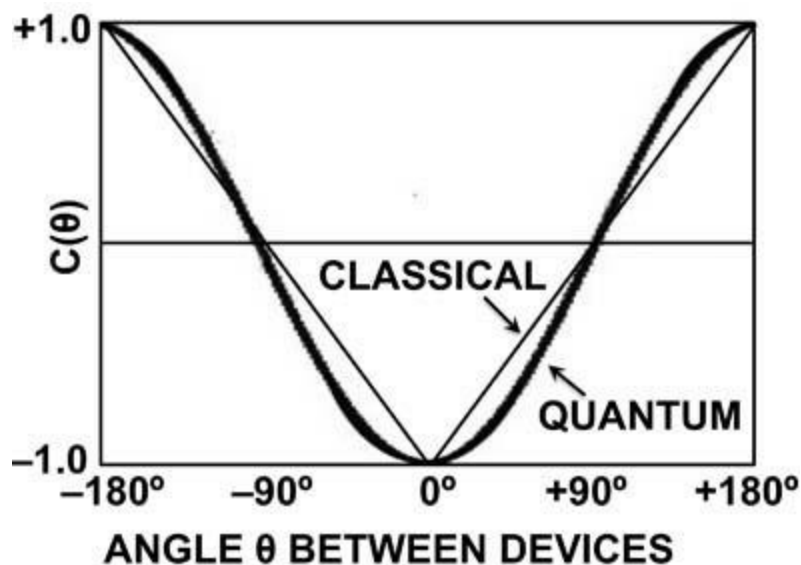


Figure 35-4 Spin Correlation vs.  $\theta$ : Classical vs. QM

From these differing predictions, Bell found an inequality limiting spin correlations in all possible classical physical theories based on Local Realism. Bell's inequality is illustrated by this example.

Suppose measurements X and Y give the same result 99% of the time (they are 99% correlated,  $C_{xy}=0.99$ ). Suppose also that Y and Z give the same result 99% of the time ( $C_{yz}=0.99$ ). According to the statistics of classical physics, X and Z must give the same result at least 98% of the time. Since X and Y differ by 1%, and Y and Z differ by 1%, X and Z can differ by no more than 2% ( $C_{xz} \geq 0.98$ ). Their correlation is stated as an inequality because X and Z could differ by less than 2%; they could in fact be identical ( $C_{xz}=1.00$  if  $X=Z$ ). The key point is that classical physics must conform to an inequality, which is:

$$(1-C_{xz}) \leq (1-C_{xy}) + (1-C_{yz})$$

We can also write this:

$$C_{xz} \geq C_{xy} + C_{yz} - 1$$

Classical physics conforms to this inequality because its spin correlation is linear in  $\theta$ :  $C(\theta) = -1 + 2\theta/\pi$ .

But, the spin correlation of quantum physics,  $C(\theta) = -\cos\theta$ , is non-linear. For  $\cos\theta = 0.99$ ,  $\cos(2\theta) = 0.96$ . This means X and Z could be only 96% correlated quantum mechanically, less correlated than any classical theory could explain.

Quantum mechanics violates Bell's inequality at some angles.

## TESTING BELL'S INEQUALITY

John Clauser and Stuart Freedman of the University of California, Berkeley, performed the first experimental test of EPRB in 1972. They used entangled photon pairs with complementary polarizations. They measured correlations over the full range of  $\theta$ , the angle between detector orientations, and confirmed the prediction of quantum mechanics at all  $\theta$ . They expressed the polarization correlation in terms of a quantity  $\delta$  from a different version of Bell's inequality. Their results are:

Prediction of classical physics:  $\delta \leq 0$   
Clauser measured:  $\delta = +0.050 \pm 0.008$

Clauser's measurement excludes the predictions of classical physics by over 6 standard deviations (99.999,999,8% confidence level). However, the detector orientations were fixed during each data run, and changed only between runs. Local Realists objected that pilot waves had ample time to sense the detector orientations and guide their particle partners accordingly.

A famous experiment in 1982 by French physicist Alain Aspect et. al. also used entangled photon pairs with complementary polarizations. Aspect measured polarization correlations using three orientations corresponding to  $\theta = 0, 22.5^\circ$  and  $45^\circ$ . The orientations were switched every 10 nanoseconds, substantially less than the light travel time of 40 nanoseconds between detectors. By switching detector orientations after photon pairs separate, Aspect intended to preclude possible pilot wave involvement. However, some contested this claim because Aspect's switching followed a set pattern, which "clever" pilot waves might detect. The polarization correlations were expressed by a quantity X that is peculiar to their setup. Their results are:

Classical physics:  $-1 \leq X \leq 0$   
Quantum physics:  $X = 0.112$   
Aspect measured:  $X = 0.101 \pm 0.020$

Aspect's result excludes the predictions of classical physics by 5 standard deviations (99.999,8% confidence level). His result is consistent with the quantum prediction, validating the assertion of quantum mechanics that entangled particle polarizations are not determined prior to measurement.

In 2010, another photon experiment, by German physicist Thomas Scheidl et. al., placed polarization detectors 144 kilometers apart, corresponding to 479 microseconds of light travel time. They employed a quantum device (light emitting diode and 50-50 beam splitter) to randomly select detector orientations every microsecond. The orientation options were  $\theta = 0, 22.5^\circ, 45^\circ$ , and  $67.5^\circ$ . They employed a different correlation criterion, expressed by a quantity Y. Their results are:

Classical physics:  $Y \leq 2.00$

Quantum physics:  $Y = 2.43$

Scheidl measured:  $Y = 2.37 \pm 0.02$

Quantum mechanics actually predicts  $Y = 2.828$  ( $2\sqrt{2}$ ) for an experimental apparatus whose detection and transmission efficiencies are 100%. Scheidl adjusted the QM prediction downward 14% to 2.43, due to their equipment's much lower efficiencies. The experiment's agreement with quantum physics is not good, being off by three standard deviations (0.27% confidence level for agreement). Nonetheless, the predictions of classical physics are excluded by 18 standard deviations, well beyond the pale. As an added bonus, the experimenters chose to do this experiment on the Canary Islands — why suffer to do science?

## Interpretation

One might think these experiments would settle the issue. Quantum mechanics certainly seems victorious after a 75-year battle. Ah, but the devil is in the details.

No physical mechanism has been offered to explain how Bob's spin is fixed by Alice's measurement. No EPRB experiment has ever observed anything connecting particles at greater than light speed. Absent a plausible explanation for that “magic”, the EPR debate may never end.

Additionally, several “loopholes” still exist regarding the EPRB experimental validation of quantum mechanics.

One loophole arises because photon detection efficiencies in these experiments are generally quite low: 5% to 30% typically. Some physicists question whether the detected photons fairly represent all entangled photon pairs emerging from the source. If an experiment detects a biased sample, its results might be skewed.

Another loophole relates to exotic pilot wave conspiracies. Ideally, both detectors would make their measurements simultaneously, and detector orientations would be selected independently, randomly, and simultaneously. The Scheidl experiment seems to achieve this, but some hard core Local Realists disagree. Derided by QM loyalists as “superdeterminists”, they point out that hidden variables might control the detector orientation selection in addition to the photon spins. That might conceivably make the detector orientations predetermined and coupled to the photon spins instead of being random and independent.

A final loophole questions the statistical nature of these experiments. Physical theories predict a correlation between the two particles of each entangled pair. Experiments estimate that correlation by averaging over thousands of pairs. Are we sure such averaging truly reflects the pair correlation?

Most physicists think these loopholes represent obsessive nitpicking. They are convinced quantum mechanics describes nature as it truly is: indeterminant, in precise accord with the Uncertainty Principle.

Some few still believe the only uncertainty exists in our understanding.

## Relational Quantum Mechanics

In 1994, Italian physicist Carlo Rovelli proposed an innovative alternative to conventional quantum mechanics, which he calls relational quantum mechanics (RQM).

Western science has traditionally stressed reductionism: discovery by breaking everything down into the smallest, and presumably simplest, parts — it is a “parts first” approach. Eastern philosophies are often more holistic — a “whole first” approach. Taking an entirely different approach, Rovelli believes relationships are *everything*.

In his 1996 paper titled “Relational Quantum Mechanics”, Rovelli says: “Quantum mechanics is a theory about the physical description of physical systems relative to other systems, and this is a complete description of the world.” The word “complete” is particularly compelling. (Recall that completeness was the focus of Einstein’s attack on quantum mechanics.) Rovelli believes each object in nature is no more and no less than the sum of its relationships with all other objects. He says it is meaningless to attempt to describe an object in isolation, the historical approach of Western science.

Conventional quantum mechanics makes a clear distinction between the observed and the observer. Elements of the micro-world are governed by quantum rules, while macro-world entities, including physicists and their instruments, are governed by the rules of classical physics. It is not we who are uncertain — all the fuzzy stuff happens on the “other side of the wall.” Rovelli says there is no wall — everything in nature must be considered a quantum system: subatomic particles, instruments, people, planets, stars, and even the entire universe.

Furthermore, Rovelli says observers’ descriptions of reality are valid only in relation to themselves. Einstein, in his theory of special relativity, said observers in different reference frames measure different values for basic quantities such as distance and time. Rovelli takes “relativity” to another level.

A “measurement” or “observation” must be viewed as a physical interaction between the two quantum systems: the observer X and the observed Z. That interaction produces a new quantum system XZ, which is inherently unique to its constituents. If a different observer Y had measured Z, their physical interaction could well have been quite different — Y and X could have obtained different results from observing Z. Both results would be “correct”, but each is correct only relative to its observer.

An additional consequence of RQM is that two observers can never measure exactly the same thing. When X observes Z, the state of Z is forever changed. If Y later tries to reproduce X’s measurement, Y will be measuring the new system XZ not the original system Z. This may result in a different outcome.

RQM maintains that statements like “this electron is spin up” must be replaced by “device X reports this electron is spin up.”

According to Rovelli, our quandaries of interpreting quantum mechanics arise from the invalid assumption, implicit in classical thinking, that reality exists in isolation, that objects have real properties independent of their surroundings.

## **RQM, Reality & EPR**

Rovelli has a remarkable view of the EPRB paradox. He asks: “Where is the paradox?”

Recall the EPRB formulation: particles A and B emerge from a starting point in an entangled singlet state, with equal amplitudes for spin combinations  $A+/B-$  and  $A-/B+$ . Eventually the particles will be far enough apart that they can no longer communicate with one another, even at the speed of light, within the time required to measure their spins. Thereafter, Alice measures the spin of A and Bob simultaneously measures the spin of B.

RQM states that Alice’s interaction with A creates a new quantum system: Alice-A, an entangled superposition of Alice-measured- $A+$  and Alice-measured- $A-$ . Similarly, on the other side, Bob and B become a quantum system: Bob-B, the entangled superposition of Bob-measured- $B+$  and Bob-measured- $B-$ . Since the measurements occur simultaneously at widely separated locations, neither Alice nor Bob can immediately know the result of the other’s measurement. In fact, no one can measure both particles’ spins simultaneously, nor can anyone immediately know the results of both measurements.

Only at some later time, as constrained by the speed of light, can signals from Alice-A and Bob-B combine at a common point. Since angular momentum is conserved in nature, when information from those two entangled quantum systems combine, the conclusion must be that the spin of A measured by Alice is opposite to the spin of B measured by Bob.

With RQM, there is no paradox.

Allow me to briefly digress and relate a personal story. While reading Carlo Rovelli’s excellent book ***Quantum Gravity***, I noticed a footnote at the bottom of page 217, which casually mentions that the EPR paradox is resolved by RQM. My father, Oreste Piccioni, spent countless hours pondering the EPR, and he and I had many fruitless discussions about this apparent paradox. Intrigued by the footnote, I emailed Carlo, told him of our long discussions, and requested a more detailed explanation. He graciously responded, and later decided to write a paper, published in 2008 with Matteo Smerlak, entitled “Relational EPR.” This paper presents the RQM solution at a professional, research science level. Carlo very kindly dedicated his paper to the memory of my father.

## **Peace in Our Time?**

With relational quantum mechanics, both sides of the EPR debate can claim victory. The predictions of quantum mechanics, which violate Bell’s inequality, are confirmed and are no longer startling. All interactions are local, and no signals travel faster than light. If we accept that reality is relative, following a tradition initiated by Einstein, quantum mechanics and Local Realism can peacefully

coexist.

Are Einstein and Bohr both smiling?

I fear that remains uncertain.

## Chapter 35 Review: Key Ideas

- Albert Einstein, Boris Podolsky, and Nathan Rosen, published the ***EPR* paper** saying a comprehensive physical theory must address everything that is *real*. They said a physical quantity must be considered *real* if it can be predicted with absolute certainty, without disturbing the system of which it is a part. They claimed quantum mechanics, with its Uncertainty Principle, fails this requirement.

EPR proposed a thought experiment in which the properties of one entangled particle depend on how its remote partner is measured. To Einstein, this was irrational. This dilemma is called the *EPR paradox*.

EPR claimed there are only two possible resolutions of this paradox: (1) a violation of the principle of special relativity that nothing can move through space faster than the speed of light; or (2) a violation of the Uncertainty Principle of quantum mechanics.

- **Locality** denies action-at-a-distance, particularly instantaneous action-at-a-distance: particles, energy, forces, and information cannot move faster than the speed of light. An effect cannot precede the arrival of a signal traveling at light speed from its cause.

**Realism** demands that physical quantities, such as a particle's energy, are real and precisely determined, independent of their interaction with anything else, including our instruments.

**Local Realism** is assumed by classical (pre-quantum) science.

EPR created an existential confrontation between the two great pillars of 20th century physics: relativity and quantum mechanics.

- How can physicists determine whether Local Realism or quantum mechanics best describes nature?

In the EPRB formulation, Local Realism says when two photons in a singlet state, emerge in opposite directions from point Q, each always has a fixed well-defined spin, and the two particles always have opposite spins. Conversely, quantum mechanics says the photons are entangled in a superposition state, in which neither has a well-defined spin before a measurement, but their total sum is always definitively zero.

In both theories, the photon spins are always measured to be exactly opposite one another. How can we distinguish between two theories that predict the same outcome? How can we know the true state of the photon spins before measuring them?

- John Stewart Bell found the path forward, devising ***Bell's Inequality Test*** to prove whether or not particle spins are well-determined before they are measured. His clever idea is measuring particle spin correlations  $C(\theta)$  as a function of the angle between detector orientations. The correlation predictions of classical and quantum physics are different. For photons these are:

$C(\theta) = -1 + 2\theta/\pi$ , for classical physics

$C(\theta) = -\cos\theta$ , for quantum mechanics

Quantum mechanics predicts less correlation at some angles than any classical theory could explain.

Experiments over the last 40 years confirm quantum mechanics convincingly, at least in the opinions of most physicists.

- Carlo Rovelli proposes an entirely different world-view in his theory of **Relational Quantum Mechanics**. He says relationships are *everything*. Rovelli believes each object in nature is no more and no less than the sum of its relationships with all other objects. Rovelli says: “Quantum mechanics is a theory about the physical description of physical systems relative to other systems, and this is a complete description of the world.” Rovelli says it is meaningless to attempt to describe an object in isolation, the historical approach of Western science.

# Chapter 36

## Review of QM Part Three

---

- In all processes driven by either the strong force or the electromagnetic force, the total number of fermions minus antifermions of each of the twelve types of elementary fermions is always conserved.

The number of bosons is not conserved, neither within one type nor summed over all types.

Only the weak force is known to change one type of elementary fermion into another. The weak force can change some types of quarks into others, but it does not change the total number of quarks. The weak force can also change some types of leptons into others, but it does not change the total number of leptons.

There are four kaons. The  $K^+$  and  $K^-$  are antiparticles of one another. The  $K^0$  and  $\bar{K}^0$  (“K-zero-bar”) contain opposite quarks, making them different particles and antiparticles of one another. The  $K^0$  and  $\bar{K}^0$  are not two independent particles, but are the basis states of a decaying two-state system whose Hamiltonian has complex components. Absent CP-violation, the neutral kaon “stationary” states would be:

$$|K_1\rangle = (|K^0\rangle - |\bar{K}^0\rangle)/\sqrt{2}$$
$$|K_2\rangle = (|K^0\rangle + |\bar{K}^0\rangle)/\sqrt{2}$$

The mass difference  $m_{K_2} - m_{K_1}$  is  $3.5 \times 10^{-12}$  MeV. This two-state system has three time constants:

$$K_1 \text{ lifetime } t_1 = 8.954 \times 10^{-11} \text{ sec}$$
$$K_2 \text{ lifetime } t_2 = 5.116 \times 10^{-8} \text{ sec}$$
$$\text{Oscillation time } t_{\text{osc}} = 1.88 \times 10^{-10} \text{ sec}$$

A  $K_2$  beam traversing matter *regenerates* a coherent  $K_1$  beam, because normal matter preferentially absorbs  $\bar{K}^0$  over  $K^0$ . This is called the *Pais-Piccioni effect*.

Because CP-violation is observed in neutral kaon decays, the actual decay eigenstates are the short-lived neutral kaon  $|K_S\rangle$  and the long-lived neutral kaon  $|K_L\rangle$  defined as:

$$\alpha |K_S\rangle = |K_1\rangle + \varepsilon |K_2\rangle$$
$$\alpha |K_L\rangle = |K_2\rangle + \varepsilon |K_1\rangle$$

Here  $\alpha^2 = (1 + \varepsilon^* \varepsilon)$ , and  $\varepsilon$  is a complex number measured to be:

$$\varepsilon = (2.228 \pm 0.011) \times 10^{-3} \exp\{i (43.5 \pm 0.5)^\circ\}$$

In weak interactions, parity violation is a 100% effect, but CP violation is only a 0.3% effect. We still have no comprehensive explanation for the origin of this tiny asymmetry.

- **Parity operator P** acting on circularly polarized photons reverses their direction of motion and exchanges RHC and LHC polarizations.

Elementary fermions and their antiparticles have opposite parity. Bosons are their own antiparticles. The parity of a composite system equals the product of the parities of all its constituents multiplied by  $(-1)^j$ , where  $j$  is the system's total angular momentum.

- In *electric dipole radiation*, the emitting atom's parity reverses, requiring an odd-parity photon state.

In *magnetic dipole radiation*, the emitting atom's parity is unchanged, requiring an even-parity photon state.

- **Positronium** is a bound system of an electron and an antielectron (a positron). The orbital ground state has a spin-one triplet state, and an odd parity, spin-zero singlet state. Conservation of both angular momentum and parity allow only the singlet state to decay to two photons; the triplet state must decay to three photons, making its lifetime 1000 times longer ( $1.42 \times 10^{-7}$  vs.  $1.25 \times 10^{-10}$  seconds).

- Summing angular momenta from several sources entails three key principles:

Firstly,  $M = \sum_n m_n$ .  $M$ , the z-component of angular momentum of a system, is simply the sum of the z-components of angular momenta of all the parts of that system.

Secondly, any entity's total angular momentum  $J$  is unchanged by rotation.

Thirdly, the final state with maximum total spin is the linear sum of the original states with equal coefficients.

- The electron wave function in a hydrogen atom, for quantum numbers  $n$ ,  $j$ , and  $m$ , is:

$$\begin{aligned} \psi_{n,j,m}(\theta, \phi, \rho, t) &= \alpha Y_{j,m}(\theta, \phi) F_{n,j}(\rho) \exp\{-iE_n t/\hbar\} \\ E_n &= -R_y/n^2 \\ \rho &= r / r_B = r / (\hbar^2/\mu e^2) \\ a_{k+1} &= 2a_k \{k/n-1\} / \{k(k+1) - j(j+1)\} \\ F_{n,j}(\rho) &= \exp\{-\rho/n\} \sum_k a_k \rho^{k+1} \end{aligned}$$

Here  $\alpha$  is the normalization factor,  $\mu$  is the electron mass, and the sum over  $k$  goes from  $k=j+1$  to  $k=n$ .

Each state with a unique set of quantum numbers  $n$ ,  $j$ , and  $m$ , may contain up to two electrons, one spin up and the other spin down. Each group of  $|n,j\rangle$  states has  $2j+1$  different  $m$ -states that may contain up to  $4j+2$  electrons. The maximum number of electrons with principal quantum number  $n$ , the capacity

of the  $n$ -shell, equals  $2n^2$ .

- Feynman says: “The Schrödinger equation has been one of the great triumphs of physics. By providing the key to the underlying machinery of atomic structure it has given an explanation for atomic spectra, for chemistry, and for the nature of matter.”

- **Quantum operators** act on quantum states producing new states. **Algebraic operators** act on functions producing new functions. For example:

Quantum vs. Algebraic

$$\underline{H} |\psi\rangle \Leftrightarrow \underline{H} \psi(x); \underline{H} = -(\hbar^2/2m)\partial^2/\partial x^2 + V(x)$$

$$\underline{x} |\psi\rangle \Leftrightarrow \underline{x} \psi(x) = x \psi(x)$$

$$\underline{p}_x |\psi\rangle \Leftrightarrow \underline{P}_x \psi(x); \underline{P}_x = (\hbar/i) \partial/\partial x$$

$$\underline{P} = (\hbar/i) (\partial/\partial x, \partial/\partial y, \partial/\partial z)$$

$$\underline{H} = (1/2m) \underline{P} \cdot \underline{P} + V(\underline{r})$$

For any operator  $\underline{A}$ , its *Hermitian adjoint*  $\underline{Z}$  is defined by:  $Z_{jk} = (A_{kj})^*$ . If  $\underline{Z} = \underline{A}$ ,  $\underline{A}$  is called *Hermitian* or *self-adjoint*.

- If  $M$  independent **measurements** are made on quantum entities in a uniform population, each measured value must be an eigenvalue. For a quantum system with discrete eigenstates  $|j\rangle$  and eigenvalues  $E_j$ , let the number of measurements with value  $E_j$  be  $M_j$ , with  $M = \sum_j M_j$ . The average measured value is:

$$\langle E_\psi \rangle = \sum_j (M_j/M) E_j$$

For large  $M$ , each  $M_j/M$  approaches the **expectation value**  $\text{Prob}(j)$ , the probability of  $|j\rangle$ , and  $\sigma_{mj}$ , the root-mean-square deviation of  $M_j$ , equals:

$$\sigma_{mj} = \sqrt{\{ M p(1-p) \}}, \text{ with } p = \text{Prob}(j)$$

If the probability of any particular outcome is small ( $p \ll 1$ ), this becomes:

$$\sigma_{mj} = \sqrt{\{ M_j \}}$$

For any physically observable quantity  $A$  associated with a quantum operator  $\underline{A}$ , the average value of  $A$  for any state  $|\psi\rangle$  can be expressed independently of basis states:

$$\langle A_\psi \rangle = \langle \psi | \underline{A} | \psi \rangle$$

For a system with continuous, normalized states, this becomes:

$$\langle A_\psi \rangle = \iint \psi^*(x) \langle x | \underline{A} | X \rangle \psi(X) dx dX$$

- The **commutator** of operators  $\underline{A}$  and  $\underline{B}$  is:

$$[\underline{A}, \underline{B}] = \underline{A} \underline{B} - \underline{B} \underline{A}$$

Examples are:

$$\underline{x} \underline{P}_x - \underline{P}_x \underline{x} = i\hbar$$

$$\underline{x} \underline{P}_y - \underline{P}_x \underline{y} = 0$$

$$[\underline{J}_x, \underline{J}_y] = i\hbar \underline{J}_z$$

If  $[\underline{A}, \underline{B}] = 0$ , the operators *commute*.

- The **time evolution** of the average  $\langle \underline{A} \rangle$  of a large sample of measurements corresponding to operator  $\underline{A}$  is:

$$d\langle \underline{A} \rangle / dt = \langle \psi | (i/\hbar) [\underline{H}, \underline{A}] + d\underline{A}/dt | \psi \rangle$$

Examples and their corresponding classical analogs are:

$$d\langle \underline{x} \rangle / dt = \underline{p}_x / m \Leftrightarrow dx/dt = p/m = v$$

$$d\langle \underline{p}_x \rangle / dt = -dV/dx \Leftrightarrow dp/dt = ma = F = -dV/dx$$

- The probability of finding a particle anywhere must always equal 1. This means probability is conserved globally. But more importantly, **probability is conserved locally** as well.

The **probability current**  $\hat{\underline{J}}$ , and its relationship to the change of probability is:

$$\hat{\underline{J}} = \{(\psi^* \underline{P} \psi) + (\psi^* \underline{P} \psi)^*\} / 2m$$

$$\partial(\text{Prob})/\partial t + (i/\hbar) \underline{P} \cdot \hat{\underline{J}} = 0$$

The time derivative of probability plus the divergence of probability current equals zero, integrated over any volume we choose. This means probability is conserved locally, within any selected volume.

- An electron's wave function describes the amplitude of finding the electron at each location, at each moment in time. But this does not mean an electron is “smeared-out” with a smooth charge density. If searched for, an electron will be found at a specific point, with all of its charge at that point.

If an immense number of particles exist in exactly the same state, Feynman says: “ $\psi^* \psi$  *can* be interpreted as the density of particles.”

- The wave function for photons is the same as the electromagnetic vector potential  $\underline{A}$ , which we can directly measure.

The remarkable thing about low-temperature phenomena, such as superconductivity, is that they allow fermions to “pair up” forming bosons, such as Cooper pairs, that combine in immense numbers in a common state, allowing a direct measurement of their wave function.

- Superconductivity is the property of conducting electric current with *absolutely zero* resistance. From actual observations, the estimated lifetime of superconducting currents is at least 100,000 years. Theoretical estimates exceed the age of the universe.

Atomic quantum effects produce the macroscopic phenomenon of superconductivity at extremely low temperatures. This is because higher-energy states “freeze out”, and behavior is dominated by the discreteness of the remaining states.

- In the presence of magnetic fields, two types of momentum are important, each with a different role. Feynman says the same two types arise in classical physics as well, and he defines these as:

“mv-momentum”:  $\mathbf{p} = m\mathbf{v}$

“p-momentum”:  $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$

Here  $\mathbf{A}$  is the vector potential defined by:

$$\mathbf{B} = \text{curl } \mathbf{A}$$

In the Coulomb gauge, divergence  $\{\mathbf{A}\} = 0$

- **Wave function collapse** is the greatest unresolved mystery of quantum mechanics. Einstein described it as the “entirely peculiar mechanism of *action at a distance*.” Quantum entities can exist in a superposition of different momenta, spins, locations, etc., but measurements always yield single definitive values for these quantities. To reconcile quantum indeterminacy with definitive measurements, wave functions must collapse everywhere simultaneously when disturbed by any substantial external influence, including scientific observation. While wave functions are mathematical constructs without physical reality, no one understands the “magical” mechanism by which wave functions collapse.

- The mathematical formalism of quantum mechanics is one of science’s greatest triumphs. However, physicists have substantially divergent views on its interpretation. What is the world-view of quantum mechanics — what does quantum mechanics really say about the fundamental processes of nature?

- Particles are **entangled** if they exist in a common and inseparable quantum state. For example, the state of the two photons emitted by the decay of singlet-state positronium is:

$$(|1\text{RHC}, 2\text{LHC}\rangle - |1\text{LHC}, 2\text{RHC}\rangle)/\sqrt{2}$$

Here,  $|1\text{RHC}, 2\text{LHC}\rangle$  is the basis state of photon #1 being RHC and photon #2 being LHC polarized. The two photons share one common wave function; neither has its own separate wave function. Each photon has a 50% probability of RHC and 50% probability of LHC polarization. But the probability of their polarizations being opposite is always 100%. If we measure photon #1 and find it is RHC, photon #2 instantaneously becomes LHC.

- **Teleportation** is the process of recreating material objects at a remote location without physically moving those entities through space. Teleportation is possible using entangled quantum states, but it cannot reach destinations that have not previously been reached by conventional means. One entangled particle pair is consumed for every quantum state teleported.

- Einstein, Podolsky, and Rosen (**EPR**) said a comprehensive physical theory must address everything that is *real*. They said a physical quantity is *real* if it can be predicted with absolute certainty, without measuring it. They claimed quantum mechanics, with its Uncertainty Principle, fails this requirement.

Given two entangled, but separated particles, A and B, EPR claims that measuring different properties of A makes corresponding properties of B *real*. But, if the particles are too far apart to interact, how can the reality of B's properties depend on how we choose to measure A? To Einstein, this was irrational. This dilemma is called the *EPR paradox*.

EPR claimed there were only two possible resolutions of this paradox: (1) a violation of the principle of special relativity that nothing can move through space faster than the speed of light; or (2) a violation of the Uncertainty Principle of quantum mechanics.

- **Local Realism** (Locality plus Realism) is assumed by pre-quantum science.

**Locality** denies action-at-a-distance, particularly instantaneous action-at-a-distance: particles, energy, forces, and information cannot move faster than the speed of light. An effect cannot precede the arrival of a signal traveling at light speed from its cause.

**Realism** demands that physical quantities, such as a particle's energy, are real and precisely determined, independent of their interaction with anything else, including our instruments.

- In **Bohm's EPRB** formulation, two photons in a singlet state emerge in opposite directions from point Q. Local Realism says each photon always has a fixed well-defined spin, and the two photons always have opposite spins. Conversely, quantum mechanics says the photons are entangled in a superposition state in which neither has a well-defined spin before a measurement, but their total sum is always definitively zero.

In both theories, the photon spins are always measured to be exactly opposite one another. How can we distinguish between two theories that predict the same outcome? How can we know the true state of the photon spins before measuring them?

John Stewart Bell devised **Bell's Inequality Test** to prove whether or not particle spins are well-determined before they are measured. His clever idea is measuring particle spin correlations  $C(\theta)$  as a function of the angle between detector orientations. At some angles, quantum mechanics predicts less correlation than any classical theory could explain. Experiments over the last 40 years convincingly confirm quantum mechanics, at least in the opinions of most physicists.

- Carlo Rovelli's theory of **Relational Quantum Mechanics** proposes an entirely different world-

view. He says relationships are *everything*. Rovelli believes each object in nature is no more and no less than the sum of its relationships with all other objects. Rovelli says: “Quantum mechanics is a theory about the physical description of physical systems relative to other systems, and this is a complete description of the world.” Rovelli says it is meaningless to attempt to describe an object in isolation, the historical approach of Western science.

- Feynman ends his course on quantum mechanics with this advice:

“I am sorry to say... that to participate in this adventure [physical science] it is absolutely imperative that you learn quantum mechanics as soon as possible. It was our hope that in this course we would find a way to make comprehensible to you at the earliest possible moment the mysteries of this part of physics.”

# Feynman's Epilogue

---

On the last day of class, at the end of his two-year introductory physics course, Feynman gave his farewell address:

“Well, I've been talking to you for two years and now I'm going to quit. In some ways I would like to apologize, and other ways not. I hope—in fact, I know—that two or three dozen of you have been able to follow everything with great excitement, and have had a good time with it. But I also know that ‘the powers of instruction are of very little efficacy except in those happy circumstances in which they are practically superfluous.’ So, for the two or three dozen who have understood everything, may I say I have done nothing but shown you the things. For the others, if I have made you hate the subject, I'm sorry. I never taught elementary physics before, and I apologize. I just hope that I haven't caused a serious trouble to you, and that you do not leave this exciting business. I hope that someone else can teach it to you in a way that doesn't give you indigestion, and that you will find someday that, after all, it isn't as horrible as it looks.

“Finally, may I add that the main purpose of my teaching has not been to prepare you for some examination—it was not even to prepare you to serve industry or the military. I wanted most to give you some appreciation of the wonderful world and the physicist's way of looking at it, which, I believe, is a major part of the true culture of modern times. (There are probably professors of other subjects who would object, but I believe that they are completely wrong.)

“Perhaps you will not only have some appreciation of this culture; it is even possible that you may want to join in the greatest adventure that the human mind has ever begun.

In my opinion, “the powers of instruction” are neither of “little efficacy” nor “practically superfluous.”

I believe great teachers dramatically enhance students' enjoyment, enthusiasm, achievement, and success in both career and life. I cherish the memories of several great teachers who certainly enriched my life. Feynman was one of the very best.

Feynman's stated goal was to challenge and inspire those few who might one day become world-leading physicists. I describe this goal as: “No Einstein left behind.” But, I believe educators should strive to reach more than “two or three dozen” out of 200 elite students — the other 170 deserve more. While almost everything in the *Feynman Lectures* is science that every successful physicist must learn eventually, his content is too much, too complex, and too rushed for mere mortals.

My goal is to bring Feynman's gift of science to all who love science. The *Feynman Lectures* provide the greatest trove of profound insight ever presented in an introductory physics course. I hope

this treasure will inspire many bright young minds throughout the world. I hope *Feynman Simplified* reduces your “indigestion” without diminishing the brilliance and passion of one of history’s greatest physicists.



# Meet The Author

---

Congratulations and thank you for reading my book. I know your time is valuable, and I sincerely hope you enjoyed this experience.

I'd like to tell you something about myself and share some stories.

First, the obligatory bio (as if 3 “tweets”-worth can define anyone): I have a B.S. in physics from Caltech, a Ph.D. in high-energy particle physics from Stanford University, and was on the faculty of Harvard University. Now “retired,” I teach at the Osher Institutes at UCLA and CSUCI, where students honored me as “Teacher of the Year.” In between, I ran eight high-tech companies and hold patents in medical, semiconductor, and energy technologies.

My goal is to help more people appreciate and enjoy science. We all know one doesn't have to be a world-class musician to appreciate great music — all of us can do that. I believe the same is true for science — everyone can enjoy the exciting discoveries and intriguing mysteries of our universe.

I've given 400+ presentations to general audiences of all ages and backgrounds, and have written 3 printed books and 29 eBooks. My books have won national and international competitions, and are among the highest rated physics books on Amazon.com. I'm delighted that two of these recently became the 2<sup>nd</sup> and 3<sup>rd</sup> best sellers in their fields.

Richard Feynman was a friend and colleague of my father, Oreste Piccioni, so I knew him well before entering Caltech. On several occasions, Feynman drove from Pasadena to San Diego to sail on our small boat and have dinner at our home. Feynman, my father, my brother and I once went to the movies to see *“Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb.”* It was particularly poignant watching this movie next to one of the Manhattan Project's key physicists.

At Caltech I was privileged to learn physics directly from this greatest scientist of our age. I absorbed all I could. His style and enthusiasm were as important as the facts and equations. Top professors typically teach only upper-level graduate classes. But Feynman realized traditional introductory physics didn't well prepare students for modern physics. He thought even beginners should be exposed to relativity, quantum mechanics, and particles physics. So he created a whole new curriculum and personally taught freshman and sophomore physics in the academic years 1961-62 and 1962-63.

The best students thrived on a cornucopia of exciting frontier science, but many others did not. Although Caltech may be the world's most selective science school, about half its elite and eager students drowned in Feynman's class. Even a classmate, who decades later received the Nobel Prize in Physics, struggled in this class. Feynman once told me that students sometimes gave him the “stink eye” — he added: “Me thinks he didn't understand angular momentum.”

Some mundane factors made the class very tough: Feynman's book wasn't written yet; class notes

came out many weeks late; and traditional helpers (teaching assistants and upper classmen) didn't understand physics the way Feynman taught it.

But the biggest problem was that so much challenging material flew by so quickly. Like most elite scientists, Feynman's teaching mission was to inspire the one or two students who might become leading physicists of the next generation. He said in his preface that he was surprised and delighted that 10% of the class did very well.

My goal is to reach the other 90%.

It's a great shame that so many had so much difficulty with the original course — there is so much great science to enjoy. I hope to help change that and bring Feynman's genius to a wider audience.

Please let me know how I can make *Feynman Simplified* even better — contact me through my [WEBSITE](#).

While you're there, check out my other books and sign-up for my newsletters.

***Printed Books***, each top-rated by Amazon readers:

- *Everyone's Guide to Atoms, Einstein, and the Universe*
- *Can Life Be Merely An Accident?*
- *A World Without Einstein*

***The Everyone's Guide Series of Short eBooks***

- Einstein:
  - His Struggles, and Ultimate Success, plus
  - *Special Relativity*: 3 Volumes, A to Z
  - *General Relativity*: 4 Volumes, from Introduction to Differential Topology
- *Quantum Mechanics*: 5 Volumes, from Introduction to Entanglement
- *Higgs, Bosons, & Fermions...* Introduction to Particle Physics
- Cosmology
  - *Our Universe*: 5 Volumes, everything under the Sun
  - *Our Place in the Universe*: a gentle overview
  - *Black Holes, Supernovae & More*
  - *We are Stardust*
  - *Searching for Earth 2.0*

- *Smarter Energy*
- *Timeless Atoms*
- *Science & Faith*



# Table of Contents

- [Feynman Simplified 3C: Quantum Mechanics Part Three](#)
- [Summary of QM Part One](#)
- [Summary of QM Part Two](#)
- [Chapter 25 Particle Physics:  \$\Lambda^0\$  Decay](#)
- [Chapter 26 Particle Physics: Neutral Kaons](#)
- [Chapter 27 Angular Momentum](#)
- [Chapter 28 Rotations Of Any Spin](#)
- [Chapter 29 Electron Wave Functions in Hydrogen](#)
- [Chapter 30 The Periodic Table of Elements](#)
- [Chapter 31 Quantum & Algebraic Operators](#)
- [Chapter 32 Probability & Wave Functions](#)
- [Chapter 33 Superconductivity: QM Writ Large](#)
- [Chapter 34 Collapse, Entanglement & Teleportation](#)
- [Chapter 35 EPR: QM vs. Local Realism](#)
- [Chapter 36 Review of QM Part Three](#)
- [Feynman's Epilogue](#)