

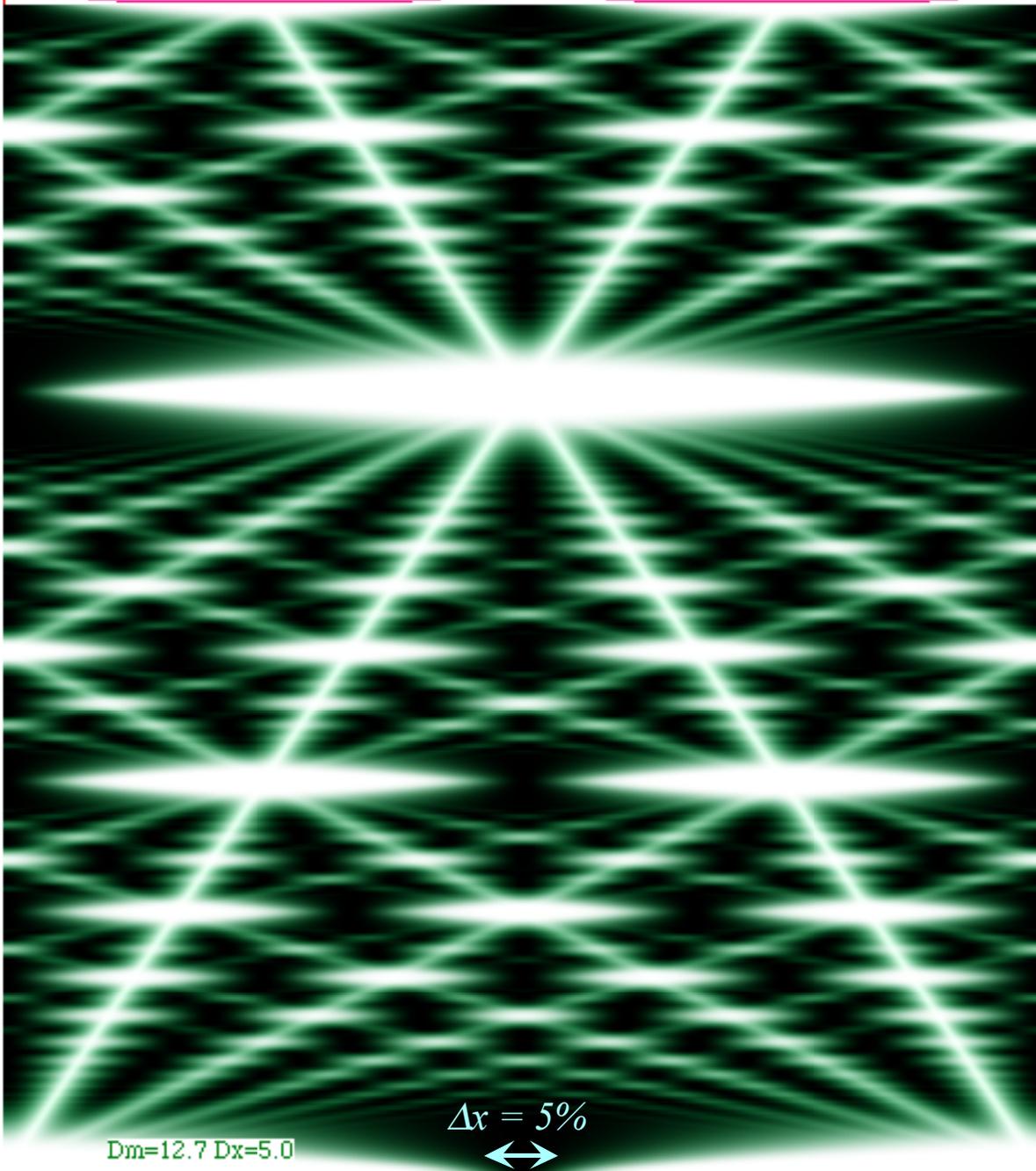
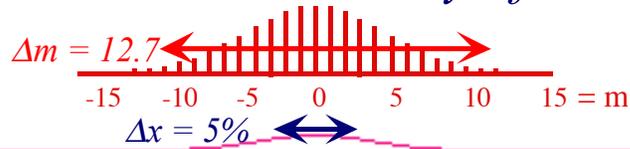
Quantum Theory for the Computer Age

W. G. Harter-University of Arkansas



Quantum Theory for the Computer Age

W. G. Harter-University of Arkansas



$Dm=12.7 Dx=5.0$

$\Delta x = 5\%$

3/4

8/11
5/7
7/10

2/3

7/11

5/8

3/5

7/12

5/9

6/11

1/2

5/11

4/9

3/7

5/12

2/5

3/8

4/11

1/3

3/10

2/7

3/11

1/4

2/9

1/5

2/11

1/6

1/7

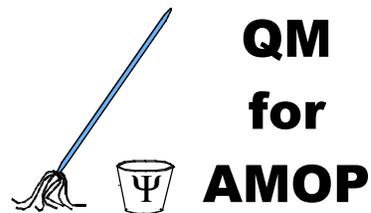
1/8

1/9

1/10
1/11
1/12

Quantum Theory for the Computer Age

An Introduction to Analysis for Atomic, Molecular, and Optical Physics



William G. Harter

Department of Physics

University
of Arkansas

Fayetteville

© 2001 thru 2013

All rights reserved

1/14/13 Draft version for Physics 5093

Hardware and Software by

HARTER-Soft

Elegant Educational Tools Since 2001

Quantum Theory for the Computer Age (QM for AMOP) W. G. Harter

Unit 1 Introduction to Quantum Amplitudes

Chapter 1 Quantum Amplitudes and Analyzers

Chapter 2 Transformation and Transfer Operators

Chapter 3 Operator Eigensolutions and Perturbations

Determinants, permanents, and permutation classes

Unit 2 Introduction to Wave Dynamics

Chapter 4 Waves Viewed by Space and Time: Relativity

Chapter 5 Waves Viewed by Wavevector and Frequency: Dispersion

Chapter 6 Multidimensional Waves and Modes

An “Old-Fashioned” classical approach to relativity

Unit 3 Introduction to Fourier Analysis and Symmetry

Chapter 7 Fourier Transformation Matrices

Chapter 8 Fourier Symmetry Analysis

Chapter 9 Time Evolution and Fourier Dynamics

Chapter 10 Two-State Evolution, Coupled Oscillation, and Spin

Optical ellipsometry and “Spin-control” using U(2) analysis

Unit 4 Introduction to Wave Equations

Chapter 11 Difference Equations and Differential Operators

Chapter 12 Infinite-Well States and Dynamics

Chapter 13 Step Potential Barriers and Wells

Scattering and “Quantum well control” using U(2) analysis

Unit 5 Introduction to Periodic Potentials and Symmetry

Chapter 14 Multiple Barriers, Eigenchannels, and Resonance Bands

Chapter 15 Non-Abelian Symmetry Analysis of Periodic Systems

Chapter 16 Fourier Analysis of Periodic Potentials and States

Molecular symmetry control inside and out

Unit 6 Introduction to Time-Variable Perturbations and Transitions

Chapter 17 Classical Electromagnetic Perturbations

Chapter 18 Transitions Due to Time-Variable Perturbation

Chapter 19 Two-State Resonant Transitions

The observer becomes the observed

Unit 7 Quantum Harmonic Oscillators

Chapter 20 One-Dimensional Oscillator States and Dynamics

Chapter 21 Two-Dimensional Oscillator States and Dynamics

Chapter 22 Quantum Electromagnetic Field

Bose-Einstein and Fermi-Dirac symmetries

Unit 8 Quantum Rotation and Angular Momentum

Chapter 23 Two-Dimensional Oscillator and Quantum Rotation

Chapter 24 Quantum Theory of Molecular and Nuclear Rotors

Chapter 25 Quantum Theory of Coupled Spins and Rotors

The quantum frame inside and out: Mach’s conundrum

Unit 9 Quantum Orbitals and Central force dynamics (In preparation)

Chapter 26 Hydrogen-like States and Dynamics

Chapter 27 Helium-like States and Dynamics

Chapter 28 Rydberg States

Fano's multichannel quantum defect theory

Unit 10 Multiparticle States and Interactions (In preparation)

Chapter 29 Unitary-Permutation Symmetry Projection

Chapter 30 $U(m) \times U(n)$ Analysis of Correlation (Entanglement)

Chapter 31 Atomic ℓ^n Configurations and Excitation

Xray photoelectron spectroscopy (XPS)

Unit 11 Polyatomic Molecules (In preparation)

Chapter 32 Molecular Orbitals and Vibration

Chapter 33 Rovibrational Fine and Superfine Structure

Chapter 34 Nuclear Hyperfine Structure

Superhyperfine spectroscopy

Unit 12 Relativistic Spin and Symmetry (Proposed)

Chapter 35 Lorentz, Poincare, and Conformal Symmetry

Chapter 36 Dirac Model of Spin

Chapter 37 Electromagnetic Interactions

Advanced potentials

Unit 13 Relativistic Quantum Field Theory (Proposed)

Chapter 35 Lorentz, Poincare, and Conformal Symmetry

Chapter 36 Dirac Model of Spin

Chapter 37 Electromagnetic Interactions

Advanced potentials

Preface

Research in quantum theory and its applications to atomic, molecular, and optical physics has grown enormously in the past half century as have related fields involving condensed matter. A new industry known as nanotechnology is just one of the results of a renaissance based on quantum mechanics. Such industry and research is largely built around an information economy, that is, computers and telecommunication.

Unfortunately, the teaching of quantum theory has not advanced as quickly. Most quantum texts follow an approach developed when slide-rules were the principle means of doing numerical computation. The notes and texts of several of the early masters, including Born, Fermi, Oppenheimer, Landau and Schwinger, form the basis of much of what current textbooks contain. This has set an orthodoxy from which few deviate.

Notable exceptions to conventional texts are ones based on lectures by Richard Feynman. The third volume of *The Feynman Lectures on Physics* by Feynman, Leighton, and Sands (*Addison Wesley 1964*) is a fresh approach to quantum theory. His unorthodox approach has survived to the present, indeed, his set of lectures are found in popular bookstores as well as in virtually every technical outlet or library in the world.

It is the Feynman approach which motivates the present work. Mentoring by Feynman and Bill Wagner (coauthor with Ferando Morinigo of *Gravitation Feynman's book*) during my introductory graduate career at Cal Tech influences this work immeasurably. The Feynman approach is characterized by an abundant use of physical analogies and pictures. While he never undersold solid mathematics, he did comment once that the disappearance of formal mathematics would “only set physics back about a week!”

One interpretation of Feynman's comment is that mathematics ought to be designed to fit the physics, not the other way around. In other words, physical insight ought to be prime mover and the main goal. Appropriate mathematics is found (or invented, if necessary) to solidify details. The result of this approach, as I hope this book shows, is better mathematics *and* physics with elegant theory, powerful computation, and most important, a set of insightful tools that uncover new directions and inventions.

One new feature of this book is something that, early on, Feynman warned against, but later he adopted fairly enthusiastically. That is the use of computer thought experiments involving both the classical commercial machines and (as yet mythical) “quantum computers.” In 1964 Feynman warned me against then new (classical) computers, “Watch out! I know guys that got sucked into those things. They're so seductive; you think you can solve anything with them!” But, by 1981 Feynman was giving lectures on computation, something he did off and on until his untimely death in 1988.

Computers play a key role in this book and one whose time has come. For over two decades I have been developing computer animations, graphics and simulations to help visualize classical and quantum phenomena. The most important outcomes of this effort have been improved physical analogies of the type that Feynman was so good at creating and did so (mostly) unaided by computers.

Times have changed, and it is difficult to say how Feynman would react to having several giga-pixel “eyes” staring back from each room in old Bridge Lab. But, we are the ones who must decide how these things are used. The approach of this book has been to make computers useful, not just for number crunching, but for the conceptual and theoretical development as well, particularly with regard to visualization of physics in space and time. Harvard educator, Howard Gardner, has noted that the human visual system has geometric pattern processing that, while less precise, handle data more quickly than its verbal or math logic processors.

An approach to quantum physics that attempts to harness a largely untapped visual human intelligence is still an academically unorthodox one. Indeed, a group of mathematicians known as Bourbakians essentially rejected figures in publications. Such images were considered childish and misleading. This small group of intellectual fundamentalists rebelled against Henri Poincare who helped create relativity and quantum theories of great and lasting value. Now after many decades, could Bourbakian claim results of greater value?

Still, it is striking that many seminal theoretical physicists including Einstein and Schwinger wrote reams of formulae without figures. In this, Feynman is a notable exception, but a prevailing attitude seems to be that analysis and calculation come first and then, perhaps, a few diagrams might be allowed.

However, if you believe as I do that physics is primarily an experimental science, then why not experiment? Whether useful ideas come from “real” lab experiments or computer lab simulations, it only matters how useful they are. One is cautious not to be misled by either one. Modern personal computers lend a graphical approach to theoretical physics that is becoming a powerful research tool. Many new basic ideas and fundamental blind spots are being exposed that might have otherwise not been found out.

Moreover, a geometric approach, such as developed in this book, should appeal to modern cyber-savvy students and enhance learning of quantum theory at all levels. We can only hope that Feynman, a pioneer in graphical physics visualization, would support this development.

About the Programs: *LearnIt* and *CodeIt*

The first tier of computer programs in this book is the *LearnIt* series consisting of *OscillIt*, *QuantIt*, *WaveIt*, etc. listed in tables below. These are (hopefully) user-friendly applications that produced many of the figures in this book. They also provide animated visualizations of physical phenomena or analogies thereof. Indeed, they are like analog computers that make text figures come alive for experimentation. Such programs were an essential aid to my ability to discover new ideas. Clearly, this needs to be made more available

The suffix “It” attached to many of these programs is derived from the *FaceIt* interface invented by Dan Kampemier, founder of *FaceWare* in Urbana, IL. It was one of the first worldwide programming projects to enhance the new Apple MacIntosh graphical user/programmer interface (GUI or GPI) and allow menus, controls, text editors, spreadsheets, movie or graphics windows to be conveniently created and used. I participated as a developer and user in *FaceWare* from 1985 until around 1993.

One advantage of *FaceWare* was that it allowed one to learn and teach *useful* root-level programming simultaneously with physics course material. The disadvantage was that it worked only on an Apple CPU and then (after Kampemier gave up) only on classic operating systems OS 7-9 or cloned OS 10.1-4. It will always be risky for research and teaching projects to develop software that relies on one type of app, GPI, or CPU. It also penalizes students who may be unable or unwilling to buy particular tool or platform.

Fortunately, there is a solution that involves high speed web browsers that are free and universally capable of running applications developed using *Java* and *HTML* programming/debugging interfaces that now exist on most of them. Dr. T. C. Reimer has pioneered in converting *LearnIt* apps from original object *Pascal*, *FORTTRAN*, and *C++* code to the modern hypertext format. Eventually, web-based text figures and formulae can become control panels for their underlying *LearnIt* applications that run seamlessly on any device.

Also, it will be possible to build a tree of programming projects for a given course that we call a *CodeIt* system. Students saw-off one or more branches of *CodeIt* trees to build their own applications as homework or lab projects. Eventually, they can build applications of sufficient complexity to aid in their thesis or dissertation research projects. Also, select *CodeIt* applications may be added to either *LearnIt* or *CodeIt* collections.

Tables below correlate the first few text chapters with some *LearnIt* programs.

Unit 1 Wave Amplitudes and Analyzers

	<i>QuantIt</i>	<i>OscillIt</i>	<i>ColorU2</i>	<i>WaveIt</i>	<i>RelativIt</i>	<i>BohrIt</i>	<i>GuidIt</i>	<i>BandIt</i>	<i>AvoidIt</i>	<i>Coullt</i>	<i>Anallt</i>
1.1	X										
1.2	X	x									
1.3	X	x	x								
1.A		x									x
1.B		x									x
2.1		x									
2.2											
3.1											
3.2											

Unit 2 Wave Dynamics (Spacetime and per-Spacetime)

	QuantIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuidIt	BandIt	AvoidIt	Coullt	Anallt
4.1			X	X							
4.2				X							
4.3				X	X						
4.4	∴	∴	∴	<u>X</u>	∴	<u>X</u>					
5.1				X	X						X
5.2				X	X	X					
5.3	X	.					
5.4	∴	∴	∴	∴	<u>X</u>	∴	∴				
6.1					X		X				
6.2					X		X				
6.3							<u>X</u>				

Unit 3 Fourier Analysis and Symmetry

	QuantIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuidIt	BandIt	AvoidIt	Coullt	Anallt
7.1						X					
7.2								X			
7.3	∴	∴	∴	<u>X</u>	∴	∴	∴	∴			
8.1				X		X					
8.2				X		X					
8.3	∴	∴	∴	<u>X</u>	∴	<u>X</u>	∴	∴			
9.1			X								
9.2											
9.3				X		X					
9.4	∴	∴	∴	<u>X</u>	∴	<u>X</u>					
10.1			X	X		X					
10.2			X	X		X		.			
10.3	∴	∴	<u>X</u>	<u>X</u>	∴	<u>X</u>	∴	∴	X		
10.4			X								

Unit 4 Wave Equations

	QuantIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuidIt	BandIt	AvoidIt	Coullt	Anallt
11.1-2				.X		X					
11.3				.		X		X			
11.4	∴	∴	∴	<u>X</u>	∴	<u>X</u>	∴	<u>X</u>			
11.5		<u>X</u>		<u>X</u>	<u>X</u>	<u>X</u>		<u>X</u>			
12.1			.			X		X			
12.2				X		X		X			
12.3			X	X		X		X			
13.1	∴	∴	∴	∴	∴	∴		X			
13.2	∴	∴	∴	∴	∴	∴		X			
13.3	∴	∴	<u>X</u>	∴	∴	∴		X			

Unit 5 Periodic Potentials

	SwingIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuideIt	BandIt	AvoidIt	Coullt	Anallt
14.1				X		X		X			
14.2				X		X		X			
15.1	:	:		<u>.X</u>	:	<u>X</u>	:	<u>X</u>			
15.2				.		.		X			
15.3				.		.		X			
15.4	:	:	:	<u>.X</u>	:	:	:	<u>X</u>			
15.5				X		X		X			
16.1	X	X	X					X			
16.2	X	X	X	.		X		X			
16.3	<u>X</u>	:	<u>.X</u>	<u>.X</u>	:	<u>X</u>		X			

Unit 6 Time Dependent Perturbation

	QuantIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuideIt	BandIt	AvoidIt	Coullt	Anallt
17.1		X									
17.2		X		.		.					
18.1				.		.					
18.2	:	:	:	:	:	:	:				
18.3				.							
19.1			X						X		

Unit 7 Quantum Harmonic Oscillators

	QuantIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuideIt	BandIt	AvoidIt	Coullt	Anallt
20.1		X				X					
20.2		X		.		<u>.X</u>					
20.3		X		.		<u>.X</u>					
21.1	:	:	<u>.X</u>	:	:	:	:				
21.2			<u>.X</u>								
21.3			X								
22.1				.							
22.2	:	:	:	:	:						

Unit 8 Oscillation, Spin, and Rotation (Under development)

	QuantIt	OscillIt	ColorU2	WaveIt	RelativIt	BohrIt	GuideIt	BandIt	AvoidIt	Coullt	Anallt
23.1	X										
23.2	X			.		.					
23.3	X			.		.					
23.4	:	:	:	:	:	:	:				
24.1			.								
24.2											
25.1				.							
25.2	:	:	:	:	:						

About the Subject Matter: A Brief Guide

This book is a *spectral* approach to quantum theory. Oscillatory phenomena including wave polarization, wave dynamics, resonance, and interference are emphasized. A student of wave optics should feel quite at home. The quantum psi-wavefunction is related in Chapter 1 to an electromagnetic E -wave field, and most waves treated in Chapters 1 through 4 relate to plane electromagnetic waves.

However, the plane waves of Chapter 4 through 6 are viewed in a new light that shows that quantum theory and relativity are quite the same subject with far simpler logic than exists in previous treatments of either one. Chapter 4 derives relativistic Doppler and Lorentz transformations by wave interference, and Chapter 5 develops relativistic matter-wave dispersion in a few simple steps. Light and matter make their own space-time coordinate manifolds by elementary spectral interference. This is a new result and one of several in this book that have only recently been published. Detailed study of elementary spectral components and their beat frequencies have shown to be a useful research tools as well as good pedagogy.

The word *spectral* has many connotations that need to be related as well as distinguished. Frequency spectra from prisms and gratings are well known *physical* phenomena since Newton, and modern (quantum) spectroscopy has increased accuracy to one part in 10^{16} or better. *Mathematical* spectra or eigenvalues of matrices often relate to laboratory spectra or quanta, but one must be careful to distinguish these two uses of the word. Chapters 1 through 3 carefully relate and distinguish physical phenomena from their mathematical descriptions, that is, distinguish physical mysteries from mathematical ones.

A key quantum concept, the *transformation matrix* $T_{ab} = \langle a|b \rangle$, is introduced in Chapter 1 and the first example is a 2-by-2 polarization rotation matrix. For a physicist, T gives outcomes of polarization experiments. Wavefunctions $\langle x|\psi \rangle = \psi(x)$ form another example of T as do wave-based Lorentz transformations in Chapter 4. But, simple 2-by-2 examples in Chapters 1-3 are quite sufficient to introduce Dirac bra-ket notation for transfer-operators describing polarization analyzers and projection-operators describing filters. “Own-states” or eigenstates of analyzer-filters are introduced as states which analyzers make or which filters pass.

Chapter 2 develops four quantum *axioms* which T -matrices obey as *physical* objects. As *mathematical* objects T -matrices relate eigenvectors of one operator to those of another. Chapter 3 connects the physical axioms to mathematical *theorems* in matrix algebra and to axioms for group algebra based on the spectral decomposition of matrices. Algebraic spectral theorems then begin to show why group algebra is powerful and fundamental to analyzing quantum spectra.

Efficient use of group algebra motivated by physics is one of the most powerful features of this book and it is introduced and explained as it is used throughout. This begins again in Chapters 7, 8, and 9 with the treatment of a “quantum-dot” system consisting of a square ($N=4$) or hexagonal ($N=6$) nano-corrall. By introducing discrete versions of Bohr’s earliest problem, an electron-on-a-ring, it is easier to introduce Fourier theory and its symmetry. Also, it corresponds to nano-devices currently being developed.

A discrete N -by- N Fourier transformation matrix made of N th roots of unity $\langle x_p|k_m \rangle = e^{imp2\pi/N}$ diagonalizes symmetry operators that satisfy $\mathbf{r}^N = \mathbf{1}$. (Such a T -matrix is known as a C_N -group character table.) At the same time the Fourier T -matrix diagonalizes all matrices that have C_N -symmetry since all such matrices are linear combinations of \mathbf{r} , \mathbf{r}^2 , \mathbf{r}^3 , ..., \mathbf{r}^N . This provides, in Chapter 8, all possible eigensolutions of all possible N -dot transfer matrices. The same is done for N -dot evolution operators or Hamiltonians $\mathbf{H} = H\mathbf{1} + S\mathbf{r} + T\mathbf{r}^2 + \dots$

in Chapter 9 and makes an elementary introduction to band theory. The approach also provides a way to introduce Schrodinger time dynamics while showing effective ways to build and solve non-trivial examples. Finally, it clarifies Bohr-matter-wave revivals and their space-time “fractal coordinates” in Chapter 9.

A general 2-by-2 Hamiltonian $\mathbf{H} = \begin{pmatrix} A & B-iC \\ B+iC & D \end{pmatrix}$ is analyzed in Chapter 10 using analogy with coupled pendulums. \mathbf{H} is expressed as a linear combination $\mathbf{H} = (A+D)/2 \sigma_0 + (A-D)/2 \sigma_A + B \sigma_B + C \sigma_C$ of reflection-symmetry operators $\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\sigma_A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\sigma_B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_C = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ to give \mathbf{H} archetypes *Type-A (Asymmetric-diagonal)* $\mathbf{H}_A = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$, *Type-B (Bilateral-balanced)* $\mathbf{H}_B = \begin{pmatrix} A & B \\ B & A \end{pmatrix} = A1 + B\sigma_B$, and *Type-C (Circular-complex-chiral-coriolis)* $\mathbf{H}_C = \begin{pmatrix} A & -iC \\ iC & A \end{pmatrix} = A1 + C\sigma_C$. Mixed types AB, AC, BC , and ABC , lead to discussion of avoided-level-crossing.

Reflection operators $(\sigma_A, \sigma_B, \sigma_C)$, apart from our pedagogical ABC -labels, are well-known Pauli spinors $(\sigma_Z, \sigma_X, \sigma_Y)$, but, it is not so well known that (apart from an i -factor) they belong to Hamilton’s *quaternion* or *hyper-complex (1, i, j, k)*-numbers found in 1841. Modern quantum theory owes a lot to this $U(2)$ algebra.

Furthermore, Hamilton’s observation of mirror-reflection properties $(\sigma^2=1)$ greatly increases their utility so they generate both quantum rotations (with the i -factor: $e^{i\sigma_a R} = 1 \cos R + i\sigma_a \sin R$) and Lorentz transformations (without the i -factor: $e^{\sigma_a L} = 1 \cosh L + \sigma_a \sinh L$). This motivates a logical development of quantum theory of spin, rotation, and relativistic wave mechanics. The ABC coupled-oscillator analogy helps make spin and quasi-spin-analogies less mysterious. In addition, there are some quite deep reasons for pursuing the coupled oscillator analogy.

An isotropic 2D-oscillator ($A=D, B=0=C$) has full $U(2)$ symmetry and so $U(2)$ leads to a much simpler theory of both quantum angular momentum and relativistic quantum field theory. Both electron spins and orbits and photon spins and orbits are simplified and unified by this in later chapters. The first $U(2)$ examples treated in Chapter 10 are photon-polarization and electron spin (as introduced in Chapter 1), and the NH_3 maser doublet. These set the stage for more advanced 2-state and N -state symmetry analysis later on.

The oldest and most prevalent (yet least studied) 2-state or $U(2)$ system is a pair of plane waves. A pair of counter-propagating plane waves are used in aforementioned Chapters 4 to 6 to derive Lorentz-Einstein relativity and quantum matter wave dispersion, two pillars of quantum theory. The $U(2)$ -wave pair system returns in Chapters 11 to 14 as a basis for analyzing eigenstates in potential barriers and wells. Using a 2-by-2 crossing or C -matrix and the scattering or S -matrix does this. The C -matrix is unimodular (as is a Lorentz matrix) while the S -matrix is unitary (as is a rotation matrix) with *eigenphase* eigenvalues $e^{i\delta}$.

The concept of *eigenchannels*, which are S -matrix eigensolutions, is developed in Chapter 13. The properties of eigenchannels and eigenphases are analyzed by ABC - $U(2)$ symmetry particularly in resonance situations where they are sensitive functions of energy and of interest for electronic-phonic-devices.

One result in Chapter 14 is an alternative to band theory in Chapter 9 that is more appropriate to treat modern super-lattice nano structures and photon band-gap devices. An important distinction is shown between *resonant* and *non-resonant* eigenchannels. The former have their largest wave amplitude *inside* a nano-structure and resemble bound state waves, while the latter pile up *outside* and resemble scattering waves. Generalization of this applies to related ebb-and-flow of molecular, atomic, nuclear, and sub-nuclear waves.

Wave symmetry analysis involving general non-Abelian (non-commutative) group theory is described in Chapter 15 using a novel approach. Again, the physical props are quantum-well or quantum-dot structures introduced before in Chapter 13 and 14. Concepts of *symmetry-relativity-duality* are introduced. These require that all transformations be defined as one wave relative to another, essentially a clarification of earlier Axioms 1-4 in Chapter 2. The result is two mutually commuting or *intertwining* groups: “outside” *global* or lab-defined symmetry operators $\{\mathbf{g}, \mathbf{g}', \dots\}$ and “inside” *local* or body-defined symmetry operators $\{\bar{\mathbf{g}}, \bar{\mathbf{g}}', \dots\}$.

The result of this extra care is an increase in computational and analytic capability with a lot simpler logic. A general Hamiltonian-matrix or S -matrix is constructed and classified in Chapter 15, as in C_N analysis of Chapter 9 or $ABC-U(2)$ analysis of Chapter 10, by its combination of symmetry operators. However, unlike Chapter 9, this symmetry is *non-commutative*, and so the matrix must be built from “inside” local operators in order to commute with all “outside” global operators. A spectral decomposition of either group leads to a related decomposition of the intertwining dual and a desired reduction of the H or S -matrix. The final result tells how much “insider” wave (like a resonant eigenchannel) and “outsider” wave (like a non-resonant eigenchannel) is present in each spectral component. The physical insight provided is considerable.

Chapter 16 rounds out the discussion of band symmetry and wave mechanics using the Fourier analysis introduced in Chapters 7 and 8. Also reintroduced are coupled pendulum models of Chapters 10 and 11 that relate Schrodinger waves in a variable potential $V(x)$ to waves along a “shower curtain” (coupled pendulums) of variable height $\ell(x)$. Momentum or k -basis representations $\langle k' | H | k \rangle$ of Hamiltonian are compared to the standard position or x -basis representation $\langle x' | H | x \rangle$. Resulting computational advantages (as well as *disadvantages*) are shown using an analogy between a space-periodic potential $V(x)$ and a time-periodic force $F(t)$ on a single pendulum. Linear resonance response is compared to multiplicative resonance or *parametric resonance*, the latter being relevant since a potential $V(x)$ acts by multiplying $\psi(x)$.

This sets up the discussion of time dependent perturbations in Chapters 17 through 19. Classical electromagnetic perturbations are described using full vector-scalar potentials $(\mathbf{A}(x,t), \Phi(x,t))$ needed to build a relativistic quantum field theory. However, the non-relativistic Schrodinger approach is developed first to satisfy prevailing electronic-photonics customs. Time is a parameter rather than a part of space-time and perturbing fields and operators are explicit functions of time governed by outside input.

Chapter 18 derives first-order perturbation theory of elementary $\mathbf{E} \cdot \mathbf{r}$ dipole resonance and Fermi-Golden-Rule constant-transition-rate theory and compares it to linear resonance of classical Lorentz theory. Chapter 19 goes beyond perturbation theory for a two-state system where the $U(2)$ -parameters $A(t)$, $B(t)$, $C(t)$, and $D(t)$ are explicit functions of time and discusses parametric resonance and Rabi NMR oscillation.

Chapters 20 to 22 develop the quantum theory of harmonic oscillation and quantum electromagnetic fields. Two-dimensional oscillator theory of Chapter 18 exploits the $U(2)$ - ABC -parameterization of Chapter 10 to begin relating $U(2)$ spin-up-spin-down to three-dimensional ABC -rotation and $R(3)$ quantum angular momentum. It also leads to *super-symmetry* since it applies to a single particle oscillating in 2D or to two particles (coupled pendulums) each oscillating in 1D. Odd oscillator quanta $n=1, 3, 5, \dots$ correspond to half-integer spin $j=1/2, 3/2, 5/2, \dots$ with odd-particle-permutation parity. Even oscillator quanta $n=0, 2, 4, \dots$ correspond to integer angular quanta $l=0, 1, 2, \dots$ with even-permutation parity. One is Bose-like the other is Fermi-like.

Chapters 23 to 25 develop the quantum theory of real $R(3)$ rotation symmetry and angular momentum using the $U(2)$ oscillator basis and Hamilton reflection symmetry. The development also uses Schwinger a - a^\dagger operator algebra and Casimir invariants. The physical props are molecular or quantum rotors that carry an intrinsic Cartesian reference frame. The full symmetry is an intertwining dual $R_{LAB}(3) * \overline{R_{BODY}(3)}$ group with global-local properties introduced in Chapter 15. As before, it leads easily to eigensolutions which here are the Wigner transformation matrices $D_{M_{Lab} N_{Body}}^{J*}$ of both half-integer- J (spinor) and integer- J (vector, tensor,...). Orbital harmonics Y^ℓ are special cases of D -functions for integer $J=\ell$: $D_{M,0}^{\ell*} = Y_M^\ell$ where the intrinsic momentum N_{Body} is set to zero and ignored. Group algebra reduces difficult issues of phase and normalization.

Group algebra also simplifies problems of coupled rotors. The most famous of these are spin-orbit (fine-structure) and spin-spin (hyperfine-structure) problems introduced in Chapter 25. Visualizing and deriving coupling transformation matrices (Clebsch-Gordan coefficients) is aided considerably by a dual-symmetry approach. This is particularly helpful for building and analyzing molecular states whose respect for various local symmetries may ebb-and-flow enormously with excitation energy.

Chapter 26 to 28 introduces atomic orbital and shell structure beginning with Coulomb orbitals that have the angular Y^ℓ -wave (derived in Chapter 23) and a radial $R_{n\ell}$ -wave. The coulomb field has an important symmetry $R(4) = R(3) \times R(3)$ that is related to the rotor symmetry of Chapter 23 and aids in calculations of eigenvalues and energy matrices. Rydberg orbitals discussed in Chapter 28 represent a large area of research in atomic spectroscopy. They are also relevant for understanding excitons in condensed matter. They should be featured as important general phenomena.

The final chapters are devoted to multiparticle systems, an enormous and ever-increasing field. Topics chosen are a tiny sampling but ones that exhibit symmetry and correlation (entanglement) effects and tools for dealing with them. The underlying symmetry of N -identical particles (molecules, nucleons, electrons, photons,...) that may occupy M quantum states is generally taken to be $U(M) \times S_N$ where S_N is the permutation symmetry of N particles. Nuclear, atomic, and molecular orbital shell theory are historically the first areas to develop this analysis. Chapter 30 and 31 introduce unitary analysis of atomic and molecular shell structure.

$U(M) \times S_N$ is part of a larger dual intertwining symmetry $U(M) \times \overline{U(N)}$ which is a most important example of the “inside*outside” quantum duality treated in Chapters 16 and 24. The $U(M)$ redefines the M -states of whichever particles they may occupy while $\overline{U(N)}$ redefines the N -particles between whatever states they may be in. The ideas of particles and states are put onto more equal and general “quasi-particle” footing. Examples are given of nuclear spins having resonantly enhanced effects on whole polyatomic molecular wavefunctions. Similar correlative effects in solids and BEC ensembles are possible.

The insight and computational power provided by these types of symmetry analyses is enormous and still largely unexplored. As quantum theory advances into the computer age, and particularly if there is to be a *quantum-computer age*, this kind of analysis is likely to advance from relative obscurity to serve its time as a methodology of quite some utility.

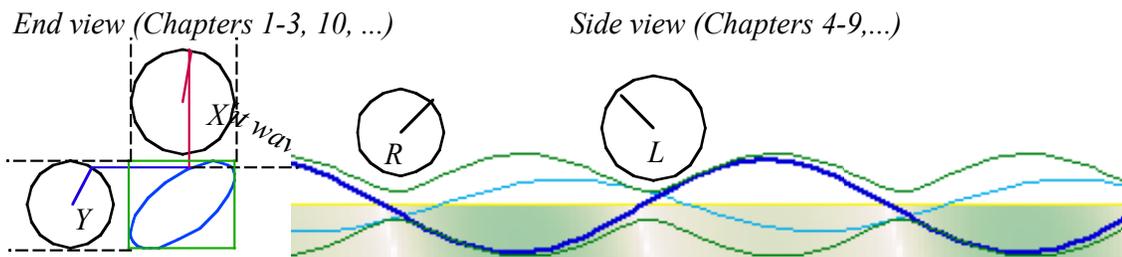
Optical Views of Quantum Mechanics

The origins of quantum theory and relativity are deeply connected with light and wave optics. Planck's axiom $E=h\nu$ was, at first, a shot in the dark, so to speak, that clarified the statistical properties of low temperature electromagnetic radiation. The history of this incredible result is found at the beginning of most texts on quantum mechanics and modern physics.

This text also uses light to develop quantum theory, but in a simpler and more direct way that avoids at first the complexity of quantum statistical mechanics. The first two units focus instead on the oscillatory wave and resonance properties of light but treat the quantum counter as a black box.

The elementary objects of thought will, for the first two units, be coherent and mostly spectrally pure laser light beams. Unit 1 concerns optical polarization, that is, light beams viewed head-on. Unit 2 concerns wave propagation, that is, light beams viewed (as best we can) from the side. In either view, (See figure below) much can be learned by modeling it as a two-state or coupled-oscillator system.

From such simple elements we develop the concept and properties of quantum matter waves by appealing to spacetime symmetry required for optical waves. It is a minimalist approach based upon analogies. It seeks to develop as much physics as possible with the simplest and least number of axioms. William Occam (1285-1349) put forth ideas known as Occam's razors to cut axioms to a minimum in order to explain the most phenomena. We hope we can use his ideas effectively in this introduction to quantum phenomena.



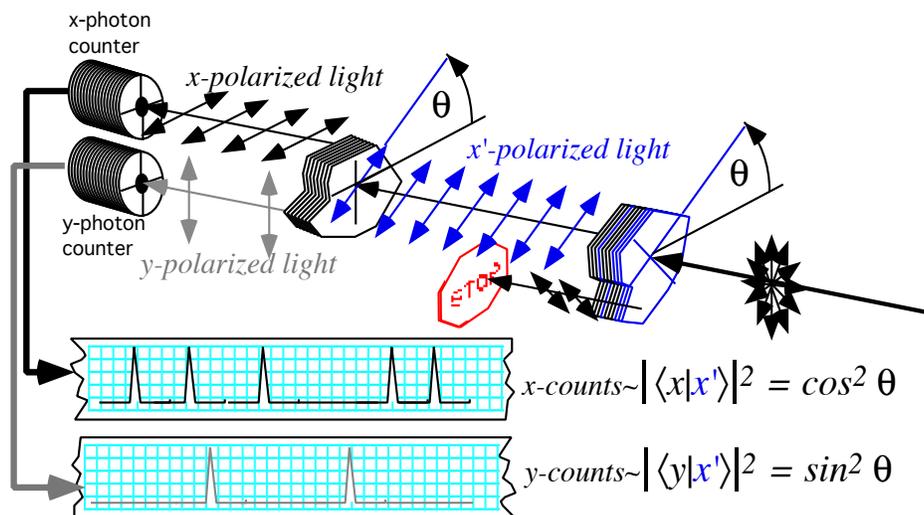


**QM
for
AMOP**

W. G. Harter

Unit 1 Quantum Amplitudes

Basic quantum amplitudes, analysis, and Dirac notation is introduced by thought experiments involving optical beams with polarization devices. Concepts such as state vectors, matrix operators, and eigensolutions are introduced via physics of beam splitters, analyzers, and counters. Operator spectral decomposition is related to state filtering and projection operators. Symmetry group operators, matrix spectral decomposition, and perturbation theory are introduced.



Chapter 1

Amplitudes, Analyzers and Matrices

The Dirac bra-ket transformation matrix $\langle a|b \rangle$ or amplitude array is introduced as the main object of study in quantum theory and related to experiments with beam sorters and analyzers. Quantum counting with and without “peeking” or dephasing is simulated and analyzed from several points of view.

UNIT 1 QUANTUM AMPLITUDES4

CHAPTER 1. AMPLITUDES, ANALYZERS AND MATRICES4

1.1 Beam Sorters.....4

 (a) Photon-beam polarization sorters.....6

 (b) Electron-beam spin polarization sorters7

1.2 Beam Sorters in Series: Transformation Matrices9

 (a) Transformation matrices for optical polarization9

 What's the state I'm in? The ideas behind projection12

 (b) WHOA! That analogy is TOO simple! Planck's energy and quantum counts13

 (c) Transformation matrices for electron spin polarization15

 (d) Amplitudes of What? Fermi vs. Bose17

 What are Photon Counters? Schrodinger's Cat (and Mouse)19

1.3 Beam Analyzers: Fundamental Quantum Processes20

 (a) Optical polarization analyzers.....20

 (1) Optical analyzers in sorter-counter configuration22

 (2) Optical analyzers in a filter configuration (Polaroid© sunglasses)23

 (3) Optical analyzers in the "control" configuration: Half or Quarter wave plates.....23

 Simulation of Active or "Do-Something" Analyzers26

 (4) Optical analyzers in a "peeking" configuration27

 (b) Effects of "peeking" : Coherent versus incoherent beams29

 (1) Amplitude products31

 (2) Amplitude sums.....31

 (3) Random phase effects ("Dirty" beams)32

 (4) Summing amplitudes or probabilities?.....33

 (c) Electron polarization analyzers34

Appendix 1.A. Review of Complex Algebra35

Appendix 1.B. Complex Response of Oscillators36

CHAPTER 2 INTRODUCTION TO TRANSFORMATION/TRANSFER OPERATORS44

2.1 Transformation Amplitude Matrices: Quantum Axioms.....45

 (a) Fundamental quantum axioms45

 (1) The probability axiom.....45

 (2) The conjugation or inversion axiom.....46

 (3) The orthonormality or identity axiom.....46

 (4) The completeness or closure axiom.....46

 (b) Matrix bra-kets: bra-and-ket vectors and representations.....47

 (1) Transformation matrix operation: Change of basis49

 (2) Transformation matrix products: Unitary groups.....49

 (4) Particle expectation number: Norms and normalization51

2.2 Transformation Operators: Unitarity and Group Axioms54

- (a) Base ket and bra transformations54
- (b) Bra-ket vector component transformations.....55
- (c) Group axioms56
 - (1) The closure axiom57
 - (2) The associativity axiom.....57
 - (3) The identity axiom.....57
 - (4) The inverse axiom57
 - (5) The commutative axiom (Abelian groups only)58
- (d) U(n) group dimension58
- (e) SU(n) group dimension58

Problems for Ch. 2.....59

CHAPTER 3 INTRODUCTION TO OPERATOR EIGENSOLUTIONS2

3.1 Operator Eigensolutions and Projection Operators.....3

Visualizing Real Symmetric Matrices and Real Eigenvectors4

- (a) Eigenvalue equations5
 - (1) Secular equations.....6
 - (2) Hamilton-Cayley equations.....7
- (b) Eigenvector projectors (Distinct eigenvalues)8
 - (1) Projector normalization9

Matrix products and eigensolutions for polarizer-counter arrangements11

 - (2) Projector completeness and spectral decomposition12

Eigensolutions are stationary or extreme-value solutions14

 - (3) Diagonalizing transformations from projectors15

Matrix products and eigensolutions for active analyzers.....16
- (c) Eigenvector projectors (Degenerate eigenvalues).....17
 - (1) Minimal equation and diagonalizability criterion17
 - (2) Nilpotent operators ("Bad" degeneracy)17
 - (3) Multiple diagonalization ("Good" degeneracy)19

Gram-Schmidt orthogonalization.....20
- (d) Projector splitting: A key to algebraic reduction20
- (e) Why symmetry groups are useful23
 - Quadratic surfaces help to visualize matrix operations24

3.2 Approximate Eigensolutions by Perturbation Techniques25

- (a) Secular determinantal expansion.....25
- (b) Perturbation approximations26
- (c) Testing perturbation approximation with exact 2x2 eigenvalues28

Appendix 3.A Matrix Determinants, Adjuncts, and Inverses1

Appendix 3.B Classification of Permutations4

UNIT. 1 REVIEW TOPICS AND FORMULAS1

Unit 1 Quantum Amplitudes

Chapter 1. Amplitudes, Analyzers and Matrices

We begin our description of quantum theory using Feynman's ideas of *particle beams* and *quantum analyzers*. A "beam-analyzer" approach lets us discuss modern atomic, molecular, and quantum optical experiments more easily than a more conventional "wavefunction-potential" approach which will be described later. Many of the newer experiments involve beams of atoms or photons which take turns undergoing "analysis." The same is true for early seminal experiments in the beginning of quantum mechanics such as those of Stern-Gerlach, Davisson-Germer, or Brown-Twiss. Our "beam-analyzer" approach will involve "thought experiments" and computer simulations based on such classic experiments.

There are theoretical reasons for using a "beam-analyzer" approach. It is more fundamental; the "wavefunction-potential" approach is a special case of the former. Also, philosophical discussion of beam-analyzer mechanics is less of a pain in the neck because many of the mysterious aspects of quantum theory are stated up-front. (In science, as in politics, a "cover-up" is usually worse than the crime.) Finally, powerful mathematical and numerical techniques are more easily motivated and understood via a "beam-analyzer" approach. This helps to demystify mathematical concepts such as operators and state vectors which might otherwise become confused with the real mystery which lies in the physics.

1.1 Beam Sorters

The fundamental idea of beam analysis is fairly simple. The basic unit is an *elementary beam sorter* which is sketched in Fig. 1.1.1. A beam sorter splits a beam of particles coming from the right into some number n of *channels*. In each channel one finds particles in some physical condition or *state* that is distinguishable from those found in neighboring channels. (The words "find" and, particularly the word "state" need to be clarified, as we will see.)

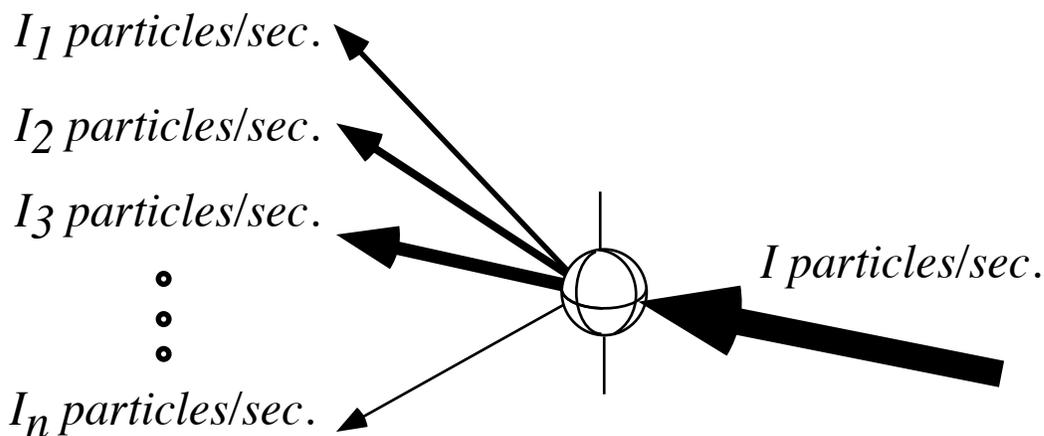


Fig. 1.1.1 Elementary beam sorter for n -state beam

Every particle that enters an elementary sorter winds up in one of the n channels ; no particles are lost or exempt. (Particles which can decay or otherwise mutate will be discussed later, but the analysis is the same; it just involves additional channels which are called *decay* or *inelastic scattering* channels.)

The initial beam (Right hand side of Fig. 1.1.1) has an *intensity* or *beam current* I . This is the number of particles per second passing a given point in that channel. This is distributed among the n channels which have currents $I_1, I_2, I_3 \dots I_n$, respectively. Particle conservation requires that these channel intensities sum up to the total I .

$$I = I_1 + I_2 + I_3 + \dots + I_n \quad (1.1.1)$$

One job of quantum mechanics is to compute *relative intensities* or *probabilities* P_k defined by

$$P_k = I_k / I \quad (1.1.2a)$$

where

$$I = P_1 + P_2 + P_3 + \dots + P_n \quad (1.1.2b)$$

follows from (1.1.1). Later, this gets "puffed up" into an operator equation called a completeness relation.

The "quantum" nature of a beam-analyzer is tacitly being assumed here. In other words, we have already begun sneaking in some pretty mysterious concepts. First, the idea of a *particle* is a quintessential quantum concept that has been (and probably will continue to be) a real mystery. It is one of those concepts that humans have taken for granted (or granite) since before the Greeks coined the word *Atmos* while observing that great stones are made of bits of sand. Perhaps, what we really mean is an *elementary* particle like an electron or a photon as opposed to a *composite* particle like Buckyball (C_{60}) or a flake of dandruff. However, that is neither a necessary nor sufficient description. For awhile, the phrase "elementary particle" was disappearing from the modern physics lexicon as it becomes increasingly clear nothing in nature is limited by our preconceived classical notion of a grain of sand. All "stuffs", meaning all forms of energy, have fundamental quantum behavior which can only mimic our preconceived notions of particles.

Second, the fact that an atomic beam can only be sorted into a finite (quantized) number n of split beams was a very big surprise when it first was observed, particularly by Stern and Gerlach whose Ag^- beam split into exactly two parts! (See Fig. 1.1.2) The curious finite splitting of beams is, perhaps, most responsible for our concept of a *quantum state*. Indeed, Goudschmitt and Uhlenbeck proposed the idea of spin-up and spin-down states of electron spin polarization to help explain a number of atomic phenomena including the Stern-Gerlach experiment.

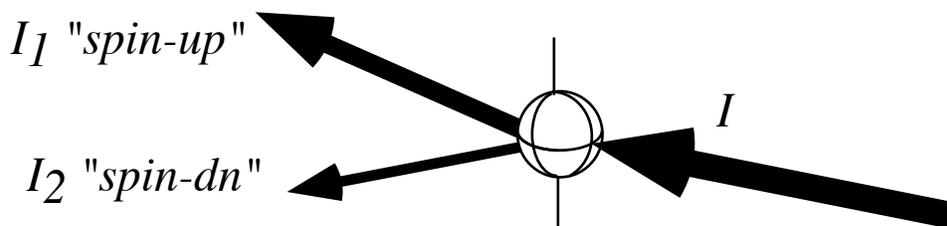


Fig. 1.1.2 Stern-Gerlach beam sorter for 2-state electron spin beam

Idealized versions of the Stern-Gerlach experiment and other two-state systems will be used to develop quantum theory in our beginning chapters. Feynman starts his description with three-state systems since their three-dimensional state-space is simple and much like the one we live in. Our choice of two-state systems is similarly motivated by the desire for simplicity and familiarity, however it uses an ultimately simpler and more fundamental analogy that goes back to 1860-1870's optical polarization theory of Poincare and Stokes.

We introduce quantum theory *vis-a-vis* photon-spin polarization, electron-spin polarization and nuclear (proton) spin-polarization; they all use similar mathematics. It also applies to NH₃ inversion-doublet states that gave us the first coherent radiation source or *maser* and marked a beginning of the laser revolution. A great deal of physics can be learned from the 2-state systems, and it also shows how to begin dealing with general *n*-state systems and much of quantum physics. Let's begin with some examples.

(a) Photon-beam polarization sorters

Consider some beam sorting experiments that a caveman could do by peering through calcite crystals. Each crystal magically gives two beams and two images, one with light polarized along the crystal's optical *x*-axis and a split-off beam having only *y*-polarized light as shown in Fig. 1.1.3 below.

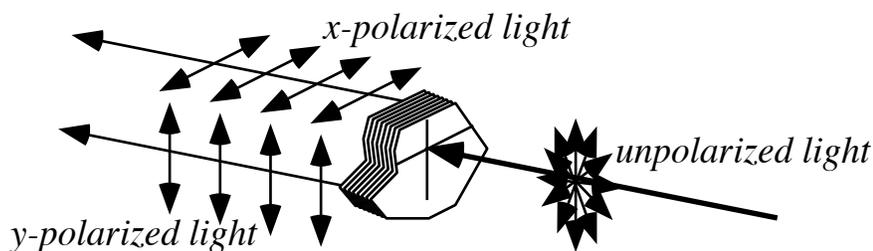


Fig. 1.1.3 Primitive photon beam sorter for 2-state polarization

If a second crystal catches the *x*-beam of the first crystal while blocking its *y*-beam, then the *y*-beam from the second crystal will disappear when the crystals' optical axes are aligned as shown in Fig. 1.1.4.

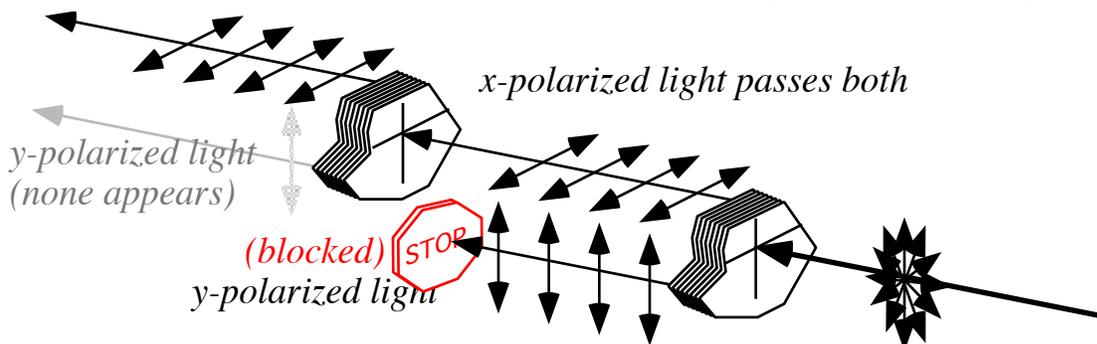


Fig. 1.1.4 Photon beam sorters in series. Second one examines *x*-beam of the first.

Modern optics labs have more sophisticated (and expensive) polarization sorters such as the Brewster prism sketched in Fig. 1.1.5. This takes advantage of fact that light reflected from a dielectric interface is nearly 100% polarized parallel to the reflection plane for a certain (Brewster's) angle of reflection.

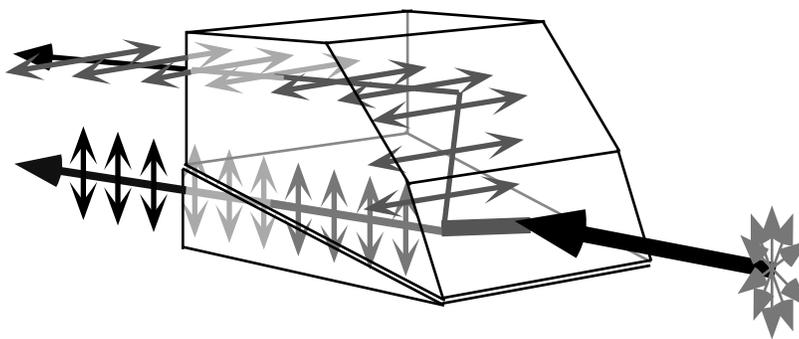


Fig. 1.1.5 Example of modern optical polarization sorter: The Brewster prism

(b) Electron-beam spin polarization sorters

Electron polarizers seem more mysterious than photon polarizers. The first ones used expensive vacuum and electron optics technology. As electronics evolved from vacuum tubes to semiconductors to *micro-m* optical fibers to *nano-m* wires, tiny *spintronic* polarizers have been developed. Here we will start with the old stuff.

A rough sketch of a Stern-Gerlach spin polarizer is shown in Fig. 1.1.6. It consists of asymmetric magnetic poles that produce a **B**-field with a large *z*-component and a field gradient tensor $\nabla\mathbf{B}$ with a large *zz*-component. The hapless electron is injected at right angles to the **B**-or *z*-axis, say, along the *y*-or beam axis. We presume electron spin angular momentum **S** and the magnetic dipole moment **m** are related by a constant of proportionality known as the *gyro magnetic ratio* γ .

$$\mathbf{S} = \gamma \mathbf{m} \tag{1.1.3}$$

A classical scenario for what happens next goes something like the following. On entry **S** and **m** are pointing more or less up-*z*-axis and moving with the electron right-to-left along the *y*-or beam axis in Fig. 1.1.6 below.

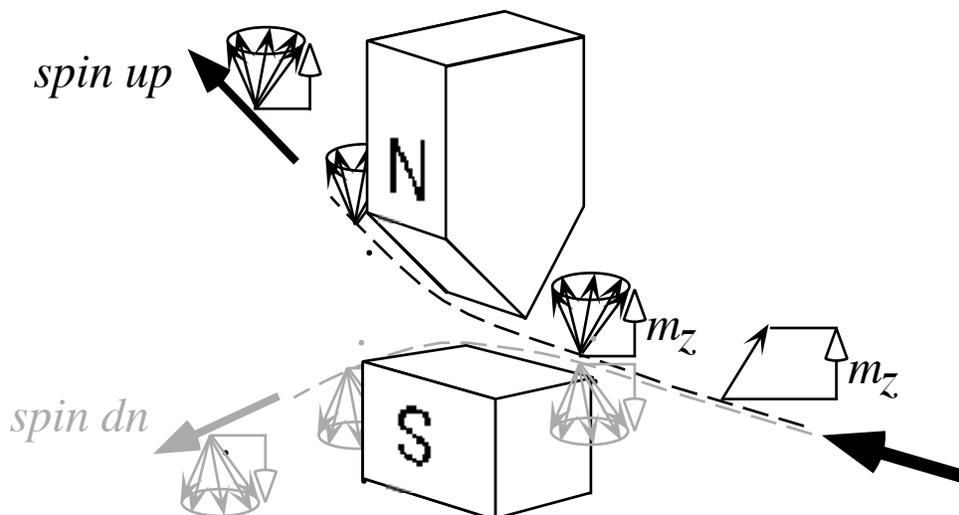


Fig. 1.1.6 Electron beam sorting by non-uniform **B**-field (Stern-Gerlach polarizer)

First, the **B**-field starts the electron spin and magnetic moment precessing like a conical helicopter blade around the **B** or *z*-axis thereby essentially freezing the *z*-component S_z or m_z of the spin moment and averaging the *x*- and *y*-components to zero. (See Fig. 1.1.6) Then the *zz*-gradient grabs the *z*-component m_z of electronic magnetic moment **m** with a force vector **F** in the direction \mathbf{e}_z of the **B**-field gradient.

$$\mathbf{F} = \mathbf{m} \bullet \nabla \mathbf{B} = \mathbf{e}_z \left(\langle m_z \rangle \frac{\partial B_z}{\partial z} \right) . \quad (1.1.4)$$

This accelerates the "helicopter" in the z -direction at a rate proportional to the z -moment component m_z that the electron had when it first encountered the \mathbf{B} -field. According to this, the final beam z -deflection is proportional to the initial z -component m_z or $S_z = \gamma m_z$ for each electron. So you might expect a randomly polarized beam to become smeared with a secant distribution up and down the left wall of the laboratory.

NOT! To practically everyone's surprise just two spots show up. The upper spot corresponds to a spin component of $S_z = +\hbar/2$ (called spin-up) and the lower spot to a spin component of $S_z = -\hbar/2$ (called spin-dn) where Planck's constant is $\hbar = h/2\pi = 1.05 \text{ E-34 Js}$. No in-between values of S_z such as zero or $\pm 0.1\hbar/2$ or $\pm 0.25\hbar/2$ are ever seen no matter how much the original beam is randomized. Each electron spin vector \mathbf{S} seems to behave like a political extremist; it chooses either to be completely up or completely down with respect to the \mathbf{B} -field. Nothing in between is ever seen. Furthermore, each electron seems to have exactly one-half quanta ($\hbar/2$) of angular momentum permanently buried in its belly. This came as a surprise to those who were just getting used to the early ideas of Bohr quantum theory which said that the smallest quantum of angular momentum or action was the Planck \hbar unit.

The Stern-Gerlach experiment also is remarkable since the electron in question is dragging along an entire silver atom which out-weighs it by a factor of about 300,000. (The experiment used a beam of Ag⁺ cations.) One could imagine dragging a 500 pound hog around by its ear!

This experiment appears to be a good deal more complicated than the cave-man polarization experiments. We shall put off discussion of its details until later, but even then, the deep-down details of electronic spin and structure remain mysterious to this day. Quantum electrodynamics (QED) has come a long way but many mysteries remain. If you can give a cogent sub-electronic theory of electron structure which explains the detailed origin of its spin 1/2 you might have a Nobel prize Fedexed to your doorstep by Friday.

For now, we can only treat spin phenomena as part of a given set of physical axioms and construct a mathematical analog for the behavior. Just such a mathematical structure is called *spinor analysis* by Jordan and Pauli who (re)discovered it around 1920. It is similar to *quaternion algebra* which was discovered by W. R. Hamilton around 1843, more than half a century before the first Stern-Gerlach experiments. The modern name for this mathematics is *U(2) group algebra* and that is one of the many mathematical ideas we will be developing. It is a credit to the efficiency of *U(2)* mathematics that it applies to both electronic and optical spin or polarization. However it does so by ignoring some obvious physical differences between these two quite diverse phenomena! We will try to address some of these as we go along.

1.2 Beam Sorters in Series: Transformation Matrices

The fun begins with investigation of beam sorters in series such as Fig. 1.2.1 below. This is the same as Fig. 1.1.4 except that the optical axis of the first crystal is tilted relative to the second by an angle θ . Without tilt ($\theta=0$) the x -polarized beam from the first crystal is 100% sorted to the x -beam exit of the second crystal and nothing shows up at the y -polarized exit as we saw in Fig. 1.1.4. But, for even a slight tilt, as in Fig. 1.2.1, there will appear a weak beam exiting the y -polarized exit of the second sorter and the x -polarized beam will be reduced in intensity by just the amount that gets diverted into the y -beam. We indicate the rotated optical axes by primed (x' , y') labels of the polarization.

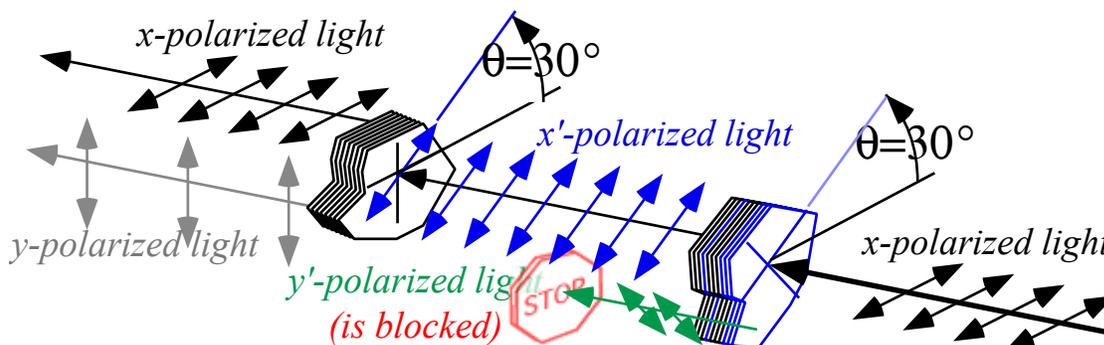


Fig. 1.2.1 Photon beam sorters in series with the first one y -blocked and tilted by angle θ .

We will use this example to introduce what is probably the single most important mathematical object in quantum theory: the *transformation matrix*. It is possible, with a little classical hand-waving, to visualize and understand the quantum transformation matrix for optical polarization. For electron polarization, which we consider subsequently, the transformation matrix and its interpretation will, at first, seem quite mysterious. Later, we will see that they are both representations of the same thing. (And, they are both mysterious, but in a nicer sort of way.)

(a) Transformation matrices for optical polarization

How does the tilted x' -polarization in Fig. 1.2.1 get transformed into y -polarized light and how much gets transformed? Consider the following model for an optically active crystal. Let it have two kinds of charged masses held by very strong springs. First there are the X -masses that can only slide and oscillate along the optical x -axis, and then there are the Y -masses which can only oscillate along the y -axis perpendicular to x . In other words, an x -polarized E-field can only wiggle the X -masses which then pass on an x -polarized polarization wave that comes out somewhere on the other side of the crystal, and similarly for y -polarized waves which come out somewhere else. In calcite the x -waves go at different speeds than the y -waves in order to produce optical beam splitting but ideally either transmits the same intensity.

Now when an electric wave with x' -polarization tilted by angle θ hits the crystal, both the X -masses and the Y -masses get stimulated in proportion to the projections $\cos \theta$ and $\sin \theta$ of the x' -direction on their respective oscillation tracks X and Y . This is sketched in Fig. 1.2.2.

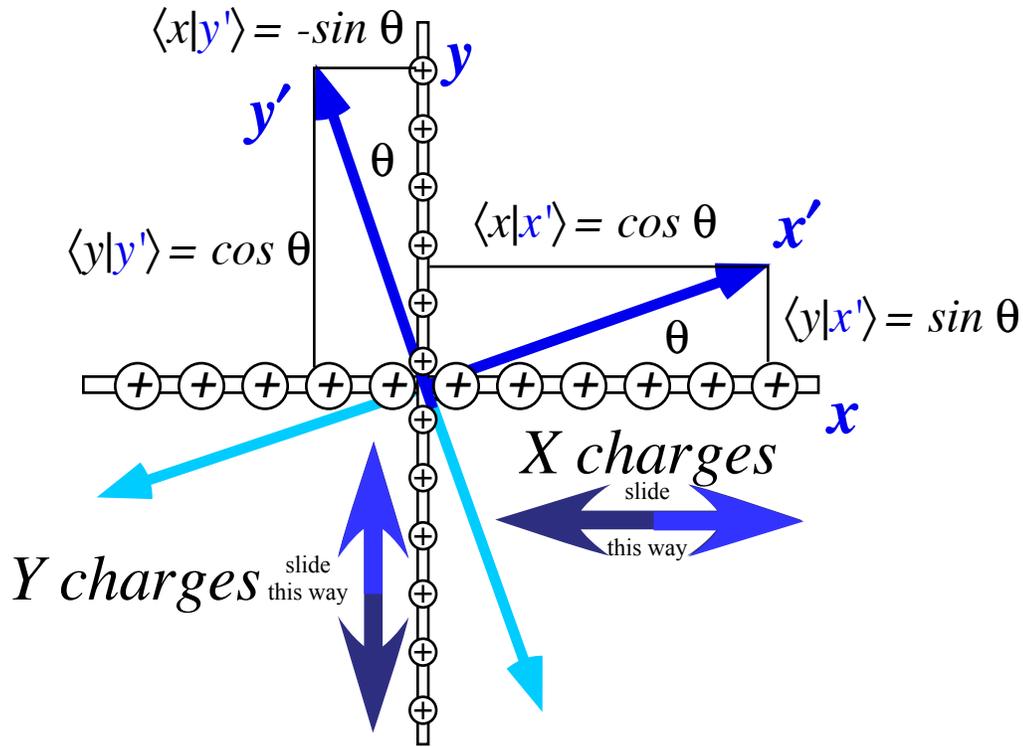


Fig. 1.2.2 Geometry of photon beam sorter for input polarizations (x',y') tilted by angle θ .

The resulting X and Y output amplitudes due to incoming x' -polarization are given by the following *Dirac bra-ket notation*.

$$\begin{aligned} X \text{ output amplitude due to } x' \text{ input} &= \langle x|x' \rangle = \cos \theta, \\ Y \text{ output amplitude due to } x' \text{ input} &= \langle y|x' \rangle = \sin \theta \end{aligned} \tag{1.2.1a}$$

If we had instead focused the y' beam (and blocked x'), then the X and Y outputs would have been the following according to Fig. 1.2.2. Note in particular that a $+y'$ field drives X-charges negatively ($-\sin\theta$).

$$\begin{aligned} X \text{ output amplitude due to } y' \text{ input} &= \langle x|y' \rangle = -\sin \theta, \\ Y \text{ output amplitude due to } y' \text{ input} &= \langle y|y' \rangle = \cos \theta \end{aligned} \tag{1.2.1b}$$

An array of these amplitudes is called the *transformation matrix* for the ideal polarization experiments of the type sketched in Fig. 1.2.1. The first column (1.2.1a) represents the x' -beam going in Fig. 1.2.1.

$$\begin{pmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \tag{1.2.1c}$$

The second column (1.2.1b) does the same for a y' -beam experiment.

The array (1.2.1c) is also a standard *coordinate transformation matrix* for rotation of coordinate axes. All quantum transformation matrices are some kind of mathematical coordinate transformation, though few are as obvious as this one. Transformation group theory is very useful in quantum mechanics.

How do you visualize and understand a transformation matrix? Dirac has given us a neat way to do so with his clever bra-ket notation. Let's take transformation (1.2.1c) apart again into its separate columns (1.2.1a) and 1.1.5b). Such columns are called *ket-vectors* or *kets* by Dirac.

$$\begin{array}{ccc}
 \begin{pmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{pmatrix} & = & \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\
 \Downarrow & & \Downarrow \\
 \begin{pmatrix} \langle x|x' \rangle \\ \langle y|x' \rangle \end{pmatrix} & = & \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad \begin{pmatrix} \langle x|y' \rangle \\ \langle y|y' \rangle \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}
 \end{array} \tag{1.2.2}$$

The kets $|x'\rangle$ and $|y'\rangle$ are just a funny notation for the unit vectors \mathbf{x}' and \mathbf{y}' indicated by the darker arrows in Fig. 1.2.2. But, their quantum mechanical significance is a bit deeper; the kets are each examples of a *polarization state vector* $|\Psi\rangle$ of a photon. The amplitudes $\langle x|\Psi\rangle$ and $\langle y|\Psi\rangle$ relate any state $|\Psi\rangle$ to the original (untilted $\theta=0$) x and y-polarization states that come out of the $\theta=0$ sorter, that is, to the basic unit vector basis $|x\rangle$ and $|y\rangle$ or \mathbf{x} and \mathbf{y} in Fig. 1.2.2 which are represented as follows.

$$|x\rangle = \begin{pmatrix} \langle x|x \rangle \\ \langle y|x \rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |y\rangle = \begin{pmatrix} \langle x|y \rangle \\ \langle y|y \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.2.3}$$

(1.2.3) is just (1.2.2) with $\theta=0$. This relation is expressed using vector sums. In Dirac notation we write

$$\begin{aligned}
 |x'\rangle &= |x\rangle\langle x|x'\rangle + |y\rangle\langle y|x'\rangle, & |y'\rangle &= |x\rangle\langle x|y'\rangle + |y\rangle\langle y|y'\rangle, \\
 &= |x\rangle(\cos\theta) + |y\rangle(\sin\theta), & &= |x\rangle(-\sin\theta) + |y\rangle(\cos\theta).
 \end{aligned} \tag{1.2.4a}$$

The same thing in Gibbs vector notation would be

$$\begin{aligned}
 \mathbf{x}' &= \mathbf{x}(\mathbf{x} \cdot \mathbf{x}') + \mathbf{y}(\mathbf{y} \cdot \mathbf{x}'), & \mathbf{y}' &= \mathbf{x}(\mathbf{x} \cdot \mathbf{y}') + \mathbf{y}(\mathbf{y} \cdot \mathbf{y}'), \\
 &= \mathbf{x}(\cos\theta) + \mathbf{y}(\sin\theta), & &= \mathbf{x}(-\sin\theta) + \mathbf{y}(\cos\theta).
 \end{aligned} \tag{1.2.4b}$$

By comparing these two notations it's clear that the transformation matrix of bra-kets corresponds to an array of dot or *scalar products*. The dot products of unit vectors are often called *direction cosines*.

$$\begin{pmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} (\mathbf{x} \cdot \mathbf{x}') & (\mathbf{x} \cdot \mathbf{y}') \\ (\mathbf{y} \cdot \mathbf{x}') & (\mathbf{y} \cdot \mathbf{y}') \end{pmatrix} \tag{1.2.5}$$

Equations (1.2.4) apply to any state $|\Psi\rangle$, not just $|x'\rangle$ or $|y'\rangle$, and to any valid quantum basis kets, not just $|x\rangle$ and $|y\rangle$. Any state can be expanded in any basis

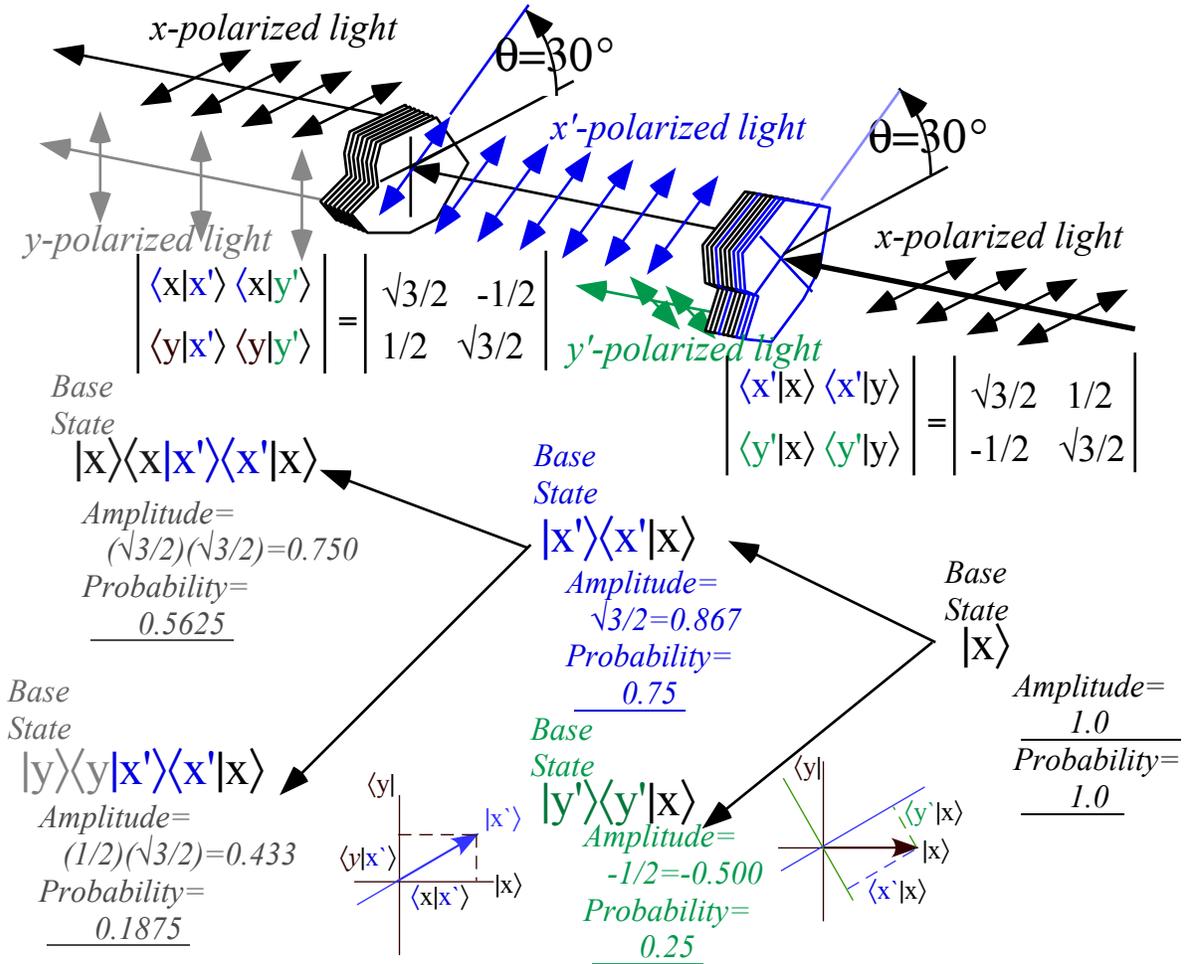
$$|\Psi\rangle = |x\rangle\langle x|\Psi\rangle + |y\rangle\langle y|\Psi\rangle = |x'\rangle\langle x'|\Psi\rangle + |y'\rangle\langle y'|\Psi\rangle \tag{1.2.6a}$$

Transformation matrices relate the amplitudes $(\langle x|\Psi\rangle, \langle y|\Psi\rangle)$ of one basis to $(\langle x'|\Psi\rangle, \langle y'|\Psi\rangle)$ of another.

$$\begin{pmatrix} \langle x|\Psi \rangle \\ \langle y|\Psi \rangle \end{pmatrix} = \begin{pmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{pmatrix} \begin{pmatrix} \langle x'|\Psi \rangle \\ \langle y'|\Psi \rangle \end{pmatrix}, \text{ or } \begin{pmatrix} \Psi_x \\ \Psi_y \end{pmatrix} = \begin{pmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{pmatrix} \begin{pmatrix} \Psi_{x'} \\ \Psi_{y'} \end{pmatrix} \tag{1.2.6b}$$

What's the state I'm in? The ideas behind projection

To get a feeling for doing quantum calculations we trace through a chain of polarization sorting experiments using the chain from Fig. 1.2.1 with a particular pure state $|x\rangle$ entering as shown below. (This is in contrast to the random mess going into Fig. 1.2.1.) We want to calculate what comes out in each channel or branch- b , namely (a) a *base state* $|b\rangle$, (b) its *amplitude* $\langle b|\Psi\rangle$, and (c) its *probability* $|\langle b|\Psi\rangle|^2$.



First is the transformation matrix $\langle b|c\rangle$ for each sorter that outputs branch b given input channel- c . The transformation matrices are given below each sorter in the figure above following (1.2.2-5).

Then the state in branch- b is $|b\rangle\langle b|\Psi\rangle$ where $|\Psi\rangle$ is whatever state came in the sorter input channel. In the x'-polarized light channel the state is simply $|x'\rangle\langle x'|x\rangle = |x'\rangle\sqrt{3}/2$. The number $\langle x'|x\rangle = \sqrt{3}/2$ is the *amplitude* of the base state of the branch or channel *base state* $|x'\rangle$, while $|\langle x'|x\rangle|^2 = 3/4$ is the *probability* or *branching ratio* for counting a particle there. This is true because any photon that makes it to branch or channel-x' must be an x'-polarized particle that is in state $|x'\rangle$ if its probability to be there is 100% or else an *attenuated* state $|x'\rangle\langle x'|x\rangle$ if its probability $|\langle x'|x\rangle|^2$ to arrive is less than one.

This process of writing $|b\rangle\langle b|\Psi\rangle$ is repeated for the next sorter in the chain only now $|\Psi\rangle$ is the previously attenuated state $|x'\rangle\langle x'|x\rangle$. So, the y-polarized branch ends up with state $|y\rangle\langle y|x'\rangle\langle x'|x\rangle$ with an even smaller amplitude $\langle y|x'\rangle\langle x'|x\rangle$ as shown in the lower left hand corner of the figure. Each use of $|b\rangle\langle b|$ is called a *projection operation* and is discussed in Sec. 2.1.(b) 5.

(b) WHOA! That analogy is TOO simple! Planck's energy and quantum counts

If simple 2D-rotation was all there was to quantum theory we probably wouldn't have courses for it! In fact, the transformation matrices, even for optical polarization, are a little more complicated than the preceding analogies might first indicate.

We mentioned that we were dealing with electromagnetic waves and charge oscillations in our simple model. We need to say a little more about this. Static (DC) polarization is fairly simple by comparison. Optical polarization involves high frequency (AC) dynamics and *resonance* phenomena. This is true for most of nature's processes, particularly those in the quantum domain where all the amplitudes wiggle incessantly like so many fidgety children. Gibb's vector notation, such as (1.2.4b) was designed for DC vectors. Dirac notation is designed for AC vectors, and AC theory uses *complex variables*.

There is more. The simple truth is this: *all quantum amplitudes are complex numbers*. At the very least they have a (sometimes hidden) time-dependent factor $e^{-i\omega t}$ given by

$$e^{-i\omega t} = \cos \omega t - i \sin \omega t \quad (1.2.7a)$$

where the angular frequency $\omega=2\pi\nu$ or frequency ν is related by *Planck's constant* $h=2\pi\hbar=6.63E-34Js$

$$\varepsilon = h\nu = \hbar\omega \quad (1.2.7b)$$

to the energy ε of a quantum state. This will be one of our most important axioms of quantum mechanics, when we get around to formal axiomization. Energy is Mother Nature's heart rate and heart beat.

In the case of light, (1.2.7b) is the equation for the energy of a single quantum of light, or *photon*, the smallest piece of energy you can extract from a light beam of a given frequency or color. Eq. (1.2.7b) is, perhaps, the first equation of quantum theory, historically and fundamentally, the basis for at least two Nobel prizes and still, many decades later, just as mysterious as it was when first stated in 1905.

However, for decades (1863-1905) classical polarization theory would ignore (1.2.7b) because the huge number of photons in a typical light beam makes it appear to be a continuous wave. The angular frequency $\omega=2\pi\nu$ of light in (1.2.7a) is presumed to be known (by color if visible) and classical resonance theory of Lorentz usually predicts polarization response due to a light beam very accurately. Generally, the beam itself was described by a complex electric field amplitude vector (E_x, E_y) which is proportional, by some constant factor f , to our unit (Ψ_x, Ψ_y) vector in (1.2.6b).

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} E_x(0)e^{-i\omega t} \\ E_y(0)e^{-i\omega t} \end{pmatrix} = f \begin{pmatrix} \Psi_x \\ \Psi_y \end{pmatrix} \quad (1.2.8)$$

Classical theory did not consider the energy $\varepsilon = \hbar\omega$ of one photon, only the *Poynting energy flux* S or *energy density* U of a whole light beam. According to Maxwell's electromagnetic wave theory the density or flux is proportional to the sum of absolute squares of the complex amplitudes.

$$S = cU, \text{ where: } U = \varepsilon_0 \left(|E_x|^2 + |E_y|^2 \right) = \varepsilon_0 \left(E_x^* E_x + E_y^* E_y \right) = \varepsilon_0 \left(E_x(0)^2 + E_y(0)^2 \right) \quad (1.2.9a)$$

Constant speed of light $c = 2.997E^8 \text{ ms}^{-1}$ and electrostatic constant $\varepsilon_0 = 8.842E-12 \text{ C}^2\text{N}^{-1}\text{m}^2$ are given.

For a beam of n -photons, the energy density is U (Joules per cubic meter) and energy flux $S=cU$ (Joules per square meter per second). Photons are assumed to have velocity c .) To relate this to that of a single photon we must equate U to n -times Planck's energy $\hbar\omega$ in eq. (1.2.7b) divided by beam or cavity volume V .)

$$\frac{n\hbar\omega}{V} = U = \epsilon_0 \left(|E_x|^2 + |E_y|^2 \right) = \epsilon_0 f^2 \left(|\Psi_x|^2 + |\Psi_y|^2 \right) \quad (1.2.10)$$

Particle number n or beam intensity I is found by dividing $U \cdot V$ by the quantum energy $\hbar\omega$.

$$n = I = \frac{V \cdot U}{\hbar\omega} = \frac{V \epsilon_0}{\hbar\omega} \left(|E_x|^2 + |E_y|^2 \right) = \frac{V \epsilon_0}{\hbar\omega} f^2 \left(|\Psi_x|^2 + |\Psi_y|^2 \right) = \frac{V \epsilon_0}{\hbar\omega} f^2 n \quad (1.2.11a)$$

For a single photon $n=I=1$ the E-field amplitude and Ψ -amplitude scale factor is given

$$f = \sqrt{\frac{\hbar\omega}{V \epsilon_0}} = \sqrt{|E_x|^2 + |E_y|^2}, \quad \text{for one-photon: } n = I = 1 \quad (1.2.11b)$$

This is the *quantum field constant* f used much later on. Note that the Ψ -amplitude squares sum to the particle number per unit time.

$$n = I = |\Psi_x|^2 + |\Psi_y|^2 = \langle x | \Psi \rangle^* \langle x | \Psi \rangle + \langle y | \Psi \rangle^* \langle y | \Psi \rangle \quad (1.2.11c)$$

This is called a *normalization condition*. For one particle ($n=I$) it is called *unit normalization*.

At first, we will set $n=I=1$, and deal with only one quantum (photon) at a time. Then each term $\Psi_x^* \Psi_x$ or $\Psi_y^* \Psi_y$ of (1.2.11c) gives the *probability* that the photon with Ψ -polarization will be found in the x -polarization state or the y -polarization state, respectively. With this statement we first confront the awful truth of quantum theory. To really do the "caveman" polarization experiment accurately we need to wait a million years or so, until the 20-th century when *photon counters* are invented. Then we buy two of these gadgets and stick them at the ends of the x and y -beams as shown in Fig. 1.2.3 below.

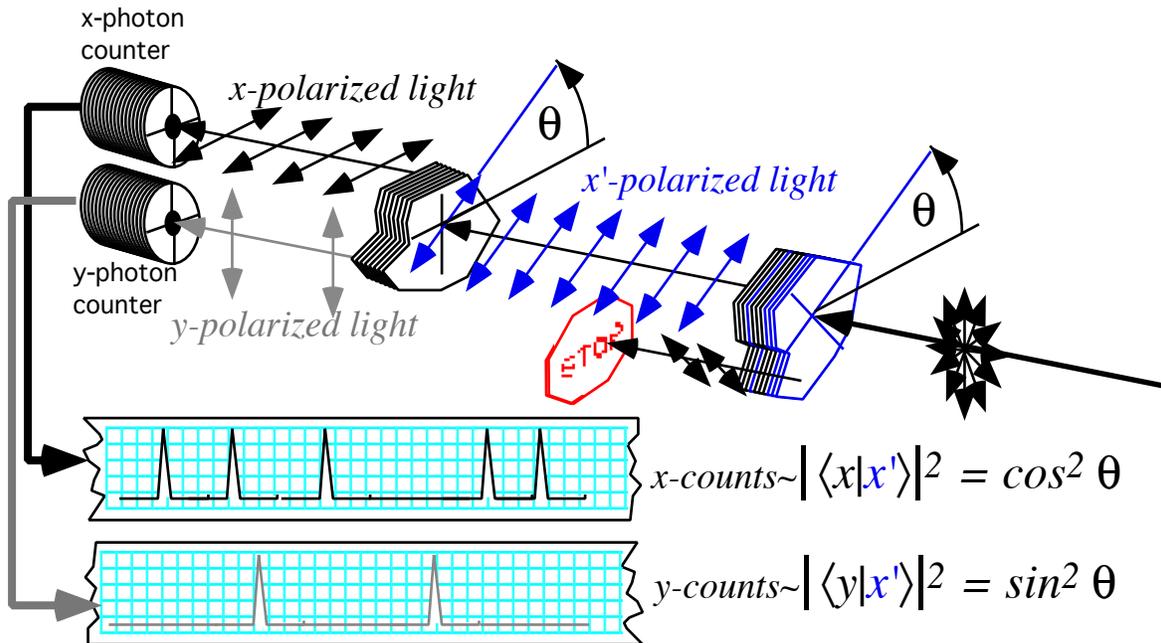


Fig. 1.2.3 Photon x - y beam sorting and quantum (photon) counting of an x' -state

After waiting for hundreds, thousands, or millions of counts the relative numbers of x -photon counts to y -photon counts will gradually approach the predicted ratios listed above in Fig. 1.2.3 of

$$|\Psi_x|^2 = \langle x | \Psi \rangle^2 = \langle x | x' \rangle^2 = \cos^2 \theta \quad \text{and} \quad |\Psi_y|^2 = \langle y | \Psi \rangle^2 = \langle y | x' \rangle^2 = \sin^2 \theta \quad (1.2.12)$$

In other words, the quantum experiment for large numbers of photons will correspond to the classical predictions of 1863. This is an example of *quantum-classical correspondence*, quantum physics usually yields classical physics in the limit of large quantum numbers or large numbers of observations.

Otherwise, quantum amplitudes yield information in the form of probabilities and statistical distributions. The absolute square $|\langle x | x' \rangle|^2$ of amplitude $\langle x | x' \rangle$ is the probability that one photon in the x' -beam will register a count in the x -counter of Fig. 1.2.3. The (complex in general) amplitude $\langle x | x' \rangle$ is called the *probability amplitude* for a x' to x transformation. We read amplitudes right to left (x' goes into x) like Hebrew because, perhaps, many of the originators of quantum theory were Jewish. Also, always remember that we square the amplitude to get the probability.

It is instructive to see some of the limitations of quantum theory early on. You might wonder, "Can quantum theory tell if a particular x' -photon will go to the x -counter or to the y -counter?" The answer appears to be a resounding NO! Not even Mother Nature, as crafty as she is, seems to know. Or you might ask, "Can we tell exactly *when* a photon will make its decision to be x or y ?" Again, NO! As we will see later, monochromatic light beams (meaning single frequency or color) are particularly reluctant to say when (or where) their individual photons are going to show up.

However, quantum theory can predict correlation statistics about the time distribution of counts, but this depends on the properties of the wizard behind the curtain on the right of Fig. 1.2.3 who is cooking up the photon beam as well as the nature of the photon counters themselves. For the time being, we will pay no attention to the wizard behind the curtain. Also, counters are assumed 100% efficient.

(c) Transformation matrices for electron spin polarization

As we said in Section 1.1.(b) electron polarization is not as easily visualized as the photon polarization. The same goes for the transformation matrices and corresponding amplitudes even though (as we will eventually see) their mathematics is virtually identical.

Indeed, the idea that an electron could and should be described as a wave was even more mysterious than the idea that light waves could be viewed as particles. Electrodynamics of the late 1800's had electrons labeled as a particles and light labeled as waves. Relativity and quantum mechanics have gone a long way toward showing the similarity of these two types of quantum energy-momentum, while also emphasizing their differences. Modern "super-unified" field theories continue attempts to unite them and all particles while modern experiments continue, more often than not, to distinguish them.

With this in mind we introduce an ideal electron polarization transformation experiment analogous to the photon polarization experiment in Fig. 1.2.3. This is shown below in Fig. 1.2.4.

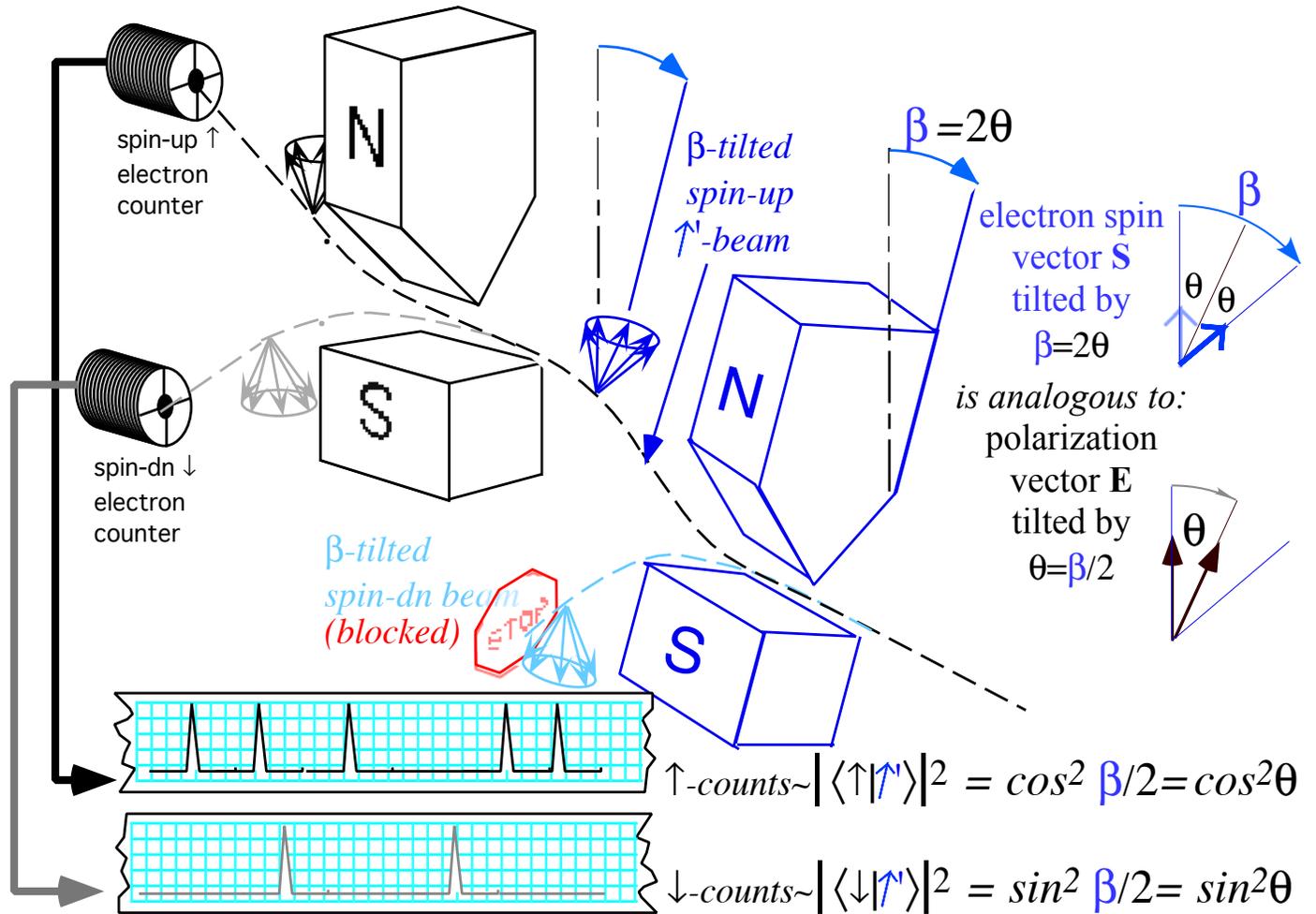


Fig. 1.2.4 Electron up-dn-spin counting of a tilted spin-up (\uparrow')-state

This is analogous to the polarization experiment first discussed after Fig. 1.2.3. Only now it is a tilted spin-up (\uparrow') electrons that have to decide whether to choose spin up (\uparrow) or spin dn (\downarrow). The up (\uparrow) and dn (\downarrow) output amplitudes due to incoming β -tilted spin-up (\uparrow') are as follows.

$$\langle \uparrow | \uparrow' \rangle = \cos \beta/2 = \cos \theta \tag{1.2.12a}$$

$$\langle \downarrow | \uparrow' \rangle = \sin \beta/2 = \sin \theta \tag{1.2.12b}$$

Comparison with (1.2.1a) shows that we use half the tilt angle ($\beta/2$) in the sine and cosine, while the photon formulas used the whole angle (θ). This is because spin-up is 180° from spin-down while x -polarization is only 90° from y -polarization. Tilting x -polarization by $\theta=90^\circ$ makes it y -polarization according to (1.2.1a) where $\cos 90^\circ=0$ and $\sin 90^\circ=1$, but tilting spin-up (\uparrow) into spin-dn (\downarrow) requires twice the angle or $\beta=2\theta=180^\circ$ according to (1.2.12a-b) where $\cos 180^\circ/2 = 0$ and $\sin 180^\circ/2 = 1$. This half-angle geometry is one of many mysterious features of the strange half-quantum spin $\hbar/2$ of an electron to be treated in later Units.

NOTE: From now on a prime means "tilted" so \uparrow' and \uparrow' mean the same tilted spin-up. Similarly for spin-dn.

An array of spin-1/2 amplitudes is called a *spinor transformation matrix*, and it describes the ideal electron spin experiments of the type sketched in Fig. 1.2.4. It is the same as the analogous photon polarization transformation matrix except the polarization angle (θ) is replaced by a half angle ($\beta/2$).

$$\begin{pmatrix} \langle \uparrow | \uparrow' \rangle & \langle \uparrow | \downarrow' \rangle \\ \langle \downarrow | \uparrow' \rangle & \langle \downarrow | \downarrow' \rangle \end{pmatrix} = \begin{pmatrix} \cos \beta / 2 & -\sin \beta / 2 \\ \sin \beta / 2 & \cos \beta / 2 \end{pmatrix} \quad (1.2.13)$$

Once again we extract columns which are called *ket-vectors* or *kets* by Dirac in analogy to (1.2.2).

$$\begin{pmatrix} \langle \uparrow | \uparrow' \rangle & \langle \uparrow | \downarrow' \rangle \\ \langle \downarrow | \uparrow' \rangle & \langle \downarrow | \downarrow' \rangle \end{pmatrix} = \begin{pmatrix} \cos \beta / 2 & -\sin \beta / 2 \\ \sin \beta / 2 & \cos \beta / 2 \end{pmatrix} \quad (1.2.14)$$

$$\begin{matrix} \downarrow & & \downarrow \\ \downarrow & & \downarrow \\ |\uparrow'\rangle = \begin{pmatrix} \langle \uparrow | \uparrow' \rangle \\ \langle \downarrow | \uparrow' \rangle \end{pmatrix} = \begin{pmatrix} \cos \beta / 2 \\ \sin \beta / 2 \end{pmatrix}, & |\downarrow'\rangle = \begin{pmatrix} \langle \uparrow | \downarrow' \rangle \\ \langle \downarrow | \downarrow' \rangle \end{pmatrix} = \begin{pmatrix} -\sin \beta / 2 \\ \cos \beta / 2 \end{pmatrix} \end{matrix}$$

The first column above represents the tilted up-(\uparrow')-beam going in to be split Fig. 1.2.1. The second column does the same for a tilted down-(\downarrow')-beam experiment. Prime (') means " β -tilted" here.

The kets are each an example of an *electron spin-state vector* $|\chi\rangle$. The amplitudes $\langle \uparrow | \chi \rangle$ and $\langle \downarrow | \chi \rangle$ relate any state $|\chi\rangle$ to the original (untilted $\beta=0$) spin-up and spin-dn states that come out of a $\beta=0$ sorter, that is, to the basic unit vector basis $|\uparrow\rangle$ and $|\downarrow\rangle$ which are represented as follows.

$$|\uparrow\rangle = \begin{pmatrix} \langle \uparrow | \uparrow \rangle \\ \langle \downarrow | \uparrow \rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} \langle \uparrow | \downarrow \rangle \\ \langle \downarrow | \downarrow \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.2.15)$$

(1.2.15) is just (1.2.14) with $\beta=0$. This relation is expressed using vector sums in Dirac notation .

$$\begin{aligned} |\uparrow'\rangle &= |\uparrow\rangle \langle \uparrow | \uparrow' \rangle + |\downarrow\rangle \langle \downarrow | \uparrow' \rangle, & |\downarrow'\rangle &= |\uparrow\rangle \langle \uparrow | \downarrow' \rangle + |\downarrow\rangle \langle \downarrow | \downarrow' \rangle, \\ &= |\uparrow\rangle \left(\cos \frac{\beta}{2} \right) + |\downarrow\rangle \left(\sin \frac{\beta}{2} \right), & &= |\uparrow\rangle \left(-\sin \frac{\beta}{2} \right) + |\downarrow\rangle \left(\cos \frac{\beta}{2} \right). \end{aligned} \quad (1.2.16)$$

Equations (1.2.16) apply to any spin state $|\chi\rangle$, not just $|\uparrow'\rangle$ or $|\downarrow'\rangle$, and to any valid quantum basis kets, not just $|\uparrow\rangle$ and $|\downarrow\rangle$. Any spin state can be expanded in any spin basis

$$|\chi\rangle = |\uparrow\rangle \langle \uparrow | \chi \rangle + |\downarrow\rangle \langle \downarrow | \chi \rangle = |\uparrow'\rangle \langle \uparrow' | \chi \rangle + |\downarrow'\rangle \langle \downarrow' | \chi \rangle \quad (1.2.17a)$$

Transformation matrices relate the amplitudes $(\langle \uparrow | \chi \rangle, \langle \downarrow | \chi \rangle)$ of one basis to $(\langle \uparrow' | \chi \rangle, \langle \downarrow' | \chi \rangle)$ of another.

$$\begin{pmatrix} \langle \uparrow | \chi \rangle \\ \langle \downarrow | \chi \rangle \end{pmatrix} = \begin{pmatrix} \langle \uparrow | \uparrow' \rangle & \langle \uparrow | \downarrow' \rangle \\ \langle \downarrow | \uparrow' \rangle & \langle \downarrow | \downarrow' \rangle \end{pmatrix} \begin{pmatrix} \langle \uparrow' | \chi \rangle \\ \langle \downarrow' | \chi \rangle \end{pmatrix}, \text{ or } \begin{pmatrix} \chi_{\uparrow} \\ \chi_{\downarrow} \end{pmatrix} = \begin{pmatrix} \langle \uparrow | \uparrow' \rangle & \langle \uparrow | \downarrow' \rangle \\ \langle \downarrow | \uparrow' \rangle & \langle \downarrow | \downarrow' \rangle \end{pmatrix} \begin{pmatrix} \chi_{\uparrow'} \\ \chi_{\downarrow'} \end{pmatrix} \quad (1.2.17b)$$

If this seems like *deja vu* (again!) from eqs. (1.2.2-6), it should. The mathematics is very similar, and we are setting the stage for general quantum theory. Electron physics, on the other hand, is quite different from photon physics when it comes to interpreting the amplitudes.

(d) Amplitudes of What? Fermi vs. Bose

As noted in Fig. 1.2.4 the electron spin-up or down count probabilities per unit time were proportional to the absolute squares of the respective amplitudes. (Actually, they're equal if you choose the right time units.)

The same was true for photon polarization x - or y counts in Fig. 1.2.3, but that was no surprise because classical field intensities have always been absolute squares of the polarization field amplitudes or wave functions.

Furthermore, the classical polarization amplitudes have had real physical meaning since about 1863 or a little before. Appropriately scaled (Recall eq. (1.2.11b).) they stood for E-fields, so many Volts per meter, or some such. At least, we thought they did!

Now along comes these post-bellum electron spin amplitudes which seem to have absolutely no physical significance whatsoever. Only their absolute squares are observable, and those are probabilities which are dimensionless apart from the (per time unit) that accompanies all particle counting experiments. There is no real physical polarization wave field analogous to the photon E-field that can be associated with an electron or even a big cloud of electrons. Why is this?

It turns out that electrons and all objects called *Fermions* which have half-quantum spins are incurable "loners." They avoid being near or like their own kind. Why this should be is still quite a mystery. Perhaps, they are embarrassed by only having half the quantum spin of the other guys and don't want anyone to find out. Whatever! So quantum physics chalks this up as one more axiom called by various names like the *Pauli principle* or *Fermi symmetry* after Wolfgang Pauli and Enrico Fermi. We will discuss these principles and their consequences in later chapters. One consequence is that electrons seem unable to get together in the same state in order to make a classical field. Fermions are tireless individualists that hold up atoms and solids but seem incapable of "unionizing" into a coherent beam.

On the other hand, photons and all objects called *Bosons* which have integral quantum spins, are real "party animals" and "copy-cat frat-rats" of the particle world and ever influenced by peer pressure. Their behavior is attributed to *Bose-Einstein symmetry* (yet another axiom) which is named after Albert Einstein and Nahari Bose. One consequence is that bosons love to swarm together into a single state and copy each other's behavior. This results in observable classical fields made of enormous numbers of bosons such those in a laser beam or atomic Bose condensate beams or liquid ^4He .

One remarkable exception is the super conductive states of electrons in various substances such as certain metals, rare-earth cuprates, and alkaline-doped solid C_{60} (buckminsterfullerite). It is thought that the electrons double up into pairs and become composite bosons which can then participate in a coherent quantum current.

To summarize: many quantum amplitudes like $\langle \uparrow | \uparrow' \rangle$ or $\langle x | \Psi \rangle$ are *probability amplitudes* only with no simple classical wave amplitude interpretation. Some see them as ghostly waves of "potential existence" or call them waves of "nothing in nothing." Only their square (probability) is real.

What are Photon Counters? Schrodinger's Cat (and Mouse)

A photon counter is like an explosive device set up to "go off" when it is disturbed by just the right stimulus. A classical analogy might be a mousetrap rigged to light a match which sets off a firecracker in a nitro factory. From miles away you can hear a big "BOOOM!!" and be pretty sure that just then a mouse nibbled its (very last) piece of cheese. Note it was the nitro factory that went "BOOOM." Mice don't go "BOOOM." By analogy, photons don't go "click" ... that's a counting device you're hearing.

Some photon counters use the *photoelectric effect* to select only photons that have an energy or frequency greater than a certain *threshold* value $E_{threshold} = h \nu_{threshold} = \hbar \omega_{threshold}$. As discussed later, an electron in an atom or molecule may be ejected from the atom in the presence of photons with energy above the *ionization threshold* for that system, but not if the photon energy is below that value. Ejected electrons are then free to be accelerated by a voltage set in the counter. This in turn ejects more electrons. After several stages of this kind of ejection and acceleration, an exponentially growing avalanche of electrons is recorded as a current "boom" or "click" that is counted. Nevertheless, photons don't go "click."

Counters have two important properties; they are amplifiers that result in *macroscopic* and, more importantly, *irreversible* effects. After the "Boom!" or the electronic avalanche, there is essentially zero probability for ever seeing an "un-Boom!" or "un-avalanche" in the lifetime of the experiment. This is quite in contrast to coherent quantum processes which allow an analyzer to sort or split up a photon and then coherently "un-sort" or reassemble the same photon at the analyzer exit as will be seen next in Fig. 1.3.1.

It is useful to describe the photon state inside the analyzer as a coherent linear combination of the form $|\Psi\rangle = \alpha|\Psi_{up}\rangle + \beta|\Psi_{dn}\rangle$. In contrast, the presence of a counter makes such a description less useful because the energy added by the amplification process will, at the very least, randomly speed up the phase of one or both of the amplitudes α or β . Moreover, it may involve coupling with huge sets of quantum states outside of the $|\Psi_{up}\rangle$ and $|\Psi_{dn}\rangle$ being studied with each outside channel siphoning its share of energy and accumulated phase and intensity. Such a complex state combination is likely to make reversibility quite impossible.

Schrodinger described a whimsical counting experiment known as *Schrodinger's Cat*. It involved a cat sleeping next to a device that poisons the poor animal only if a certain state, say $|\Psi_{up}\rangle$, is counted in the "quantum part" of the device. Schrodinger may have speculated that the state of the system could be written as a combination $\alpha|\Psi_{up}Cat_{dead}\rangle + \beta|\Psi_{dn}Cat_{alive}\rangle$, that is a cat $|\alpha|^2$ -percent-dead and $|\beta|^2$ -percent-alive.

However, this is nonsense since the experiment proposed is just a particularly complicated (and cruel) counting amplifier. The amplitudes α or β have no predictive value beyond the statistical probabilities $|\alpha|^2$ or else $|\beta|^2$ which might, under the most ideal conditions, just give approximate actuary for, say, a billion cat experiments. Such a Frankenstein laboratory cat experiment is unlikely to have any more coherence than a "Schrodinger's mouse" experiment mentioned above.

Nevertheless, the term *Schrodinger's Cat* is used to describe macroscopic and coherently reversible coupled-atom-cavity experiments that have been studied. There is no readily apparent theoretical limit to the size and complexity of a state in which quantum amplitudes like α and β may be coherently phased, and now there is evidence that quantum coherence is necessary to explain some biochemical pathways. Could actual cats (or mice) participate in coherent quantum experiments anytime in the foreseeable future? Meow!

1.3 Beam Analyzers: Fundamental Quantum Processes

To describe the weirdness and beauty of quantum physics we will discuss experiments involving elementary *quantum analyzers*. The quantum analyzer we will use here consists of two beam sorters of the type discussed in Section 1.2 placed back-to-back. More correctly we will pair up a beam sorter with a *beam un-sorter* or "put-back-together-er" that exactly undoes the splitting caused by its sorting companion.

Is this even possible? Yes, we know it is, provided you choose your apparatus judiciously and construct it carefully. One of the deepest axioms for much of atomic and molecular physics is the idea of perfect *time-reversal symmetry* at a sub-microscopic level. Roughly speaking, it says, "Anything you can do, I can do backwards." or "If you do, I can undo!" Let's see some examples of analyzer experiments.

(a) Optical polarization analyzers

Optical polarization analyzers are routinely constructed in a modern laser laboratory. Two Brewster prisms of the type sketched in Fig. 1.1.5 have only to be carefully mounted back-to-back as shown in Fig. 1.3.1. By "carefully" we mean that one must adjust split beam paths so that the output beam has exactly the same polarization and intensity as the input beam, that is, $|\Psi_{\text{OUT}}\rangle = |\Psi_{\text{IN}}\rangle$. Ideally, this would mean matching the x and y optical path lengths to within a fraction of the 0.5 micron wavelength of light. An accuracy of \pm one ten-millionths of a meter might be enough. It's definitely not a "caveman" device!

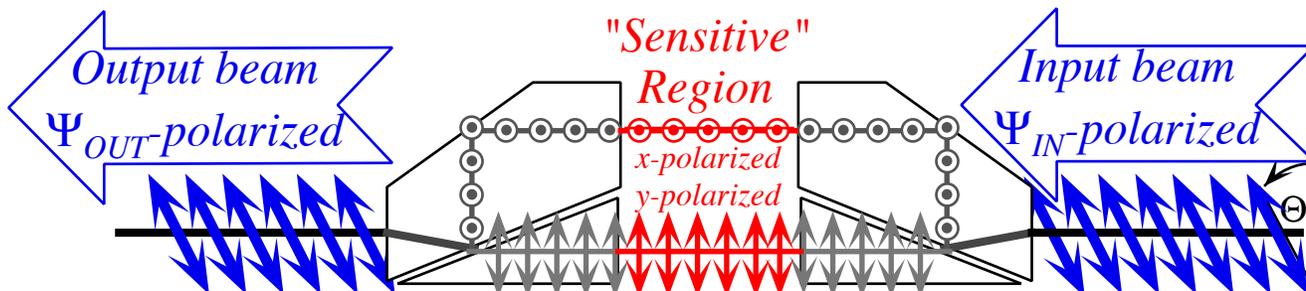


Fig. 1.3.1 Anatomy of ideal optical polarization analyzer

The purpose of these analyzers is let us to "tickle" or "perturb" each photon as it goes through a "sensitive" region where it has been temporarily sorted into to two different polarization states. Then you see what happens to its polarization as it emerges supposedly "reassembled." It is important to understand how incredibly sensitive the photon state is to what happens between the points where it is sorted on the right hand side of Fig. 1.3.1 and "un-sorted" on the left hand side. In between sorting and un-sorting, the physicists play the role of the Marquis de Sade with the poor photons in such devices. Soon the "Photon-Rights" groups may be picketing their darkened laboratories!

A theorist avoids "photon sadism" by simulating analyzer experiments on a computer. The following discussions use the *QuantIt* program which draws visual representations of the amplitudes in split beams. A *QuantIt* simulation of a "do-nothing" analyzer set up like Fig. 1.3.1 is shown in the Fig. 1.3.2 . The analyzer is shown receiving a beam of $\theta = \beta/2 = 30^\circ$ polarized photons from the right and sending out the same polarization toward the left. (It "does nothing" to polarization.) In between there is a "high road" beam that is x-polarized ($\theta = \beta/2 = 0^\circ$) and a "low road" beam that is y-polarized ($\theta = \beta/2 = 90^\circ$).

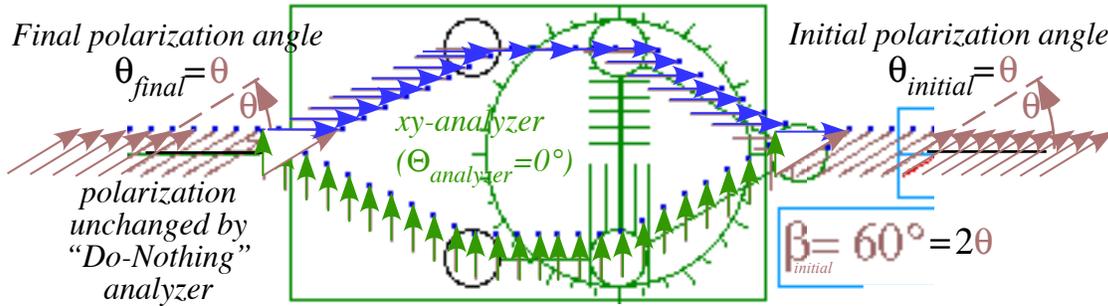
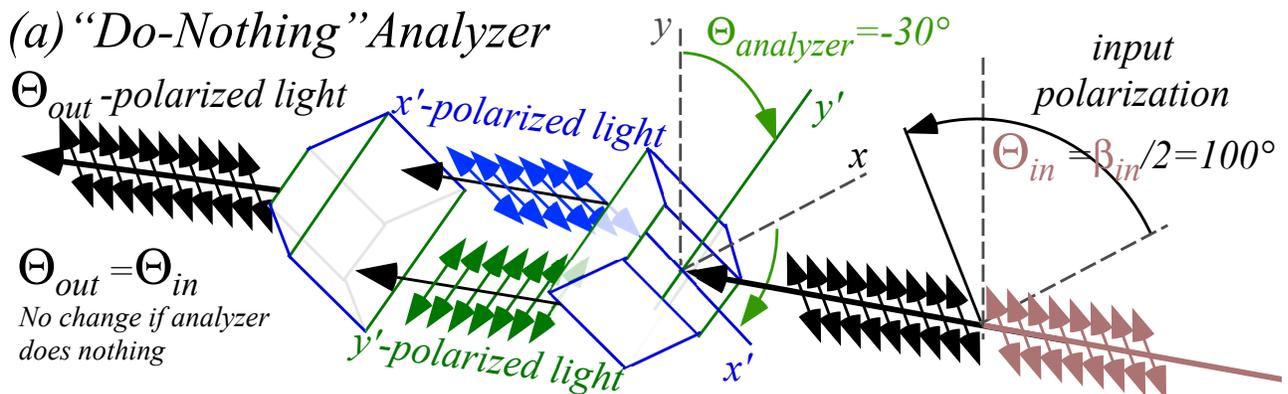


Fig. 1.3.2 Computer sketch of simulated polarization analyzer in "do-nothing" mode

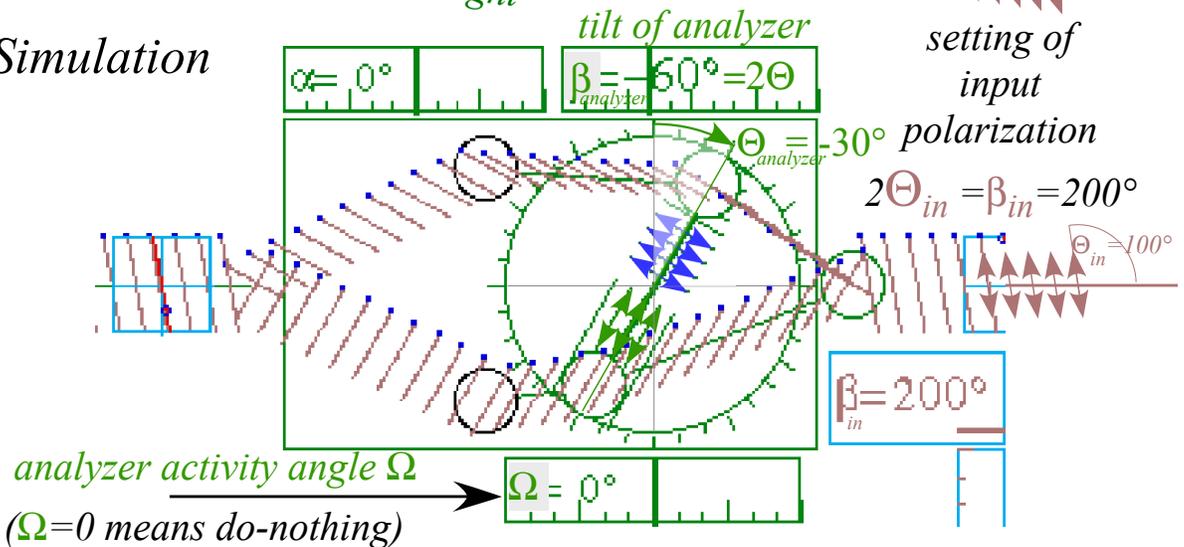
Note that polarization **E**-vectors are indicated by little lines drawn in the plane of the beam paths since it is impractical in this kind of figure to draw them as they really are (transverse to beam) while showing beam paths, too. Another "do-nothing" example is shown in the side-bar below. Following that are some examples of analyzer experiments and configurations that "do something."

Simulation of "Do-Nothing" Analyzers by QuantIt

Various analyzer configurations are simulated by *QuantIt* as shown by a do-nothing ($\Omega=0^\circ$) analyzer in the figure below. Initially polarized state enters from the right, perhaps from another analyzer that is set initially to input angle β_{in} . (Here it is set initially to $\Theta_{in} = 100^\circ$ or $\beta_{in} = 200^\circ$.) *QuantIt* uses electron spin-tilt- β -angles that double photon-x-polarization tilt Θ -angles: $\beta=2\Theta$. (Recall electron vs. photon analogy in Fig. 1.2.4.)



(b) Simulation



To calculate Θ -analyzer output imagine another analyzer, say, a basic xy -analyzer, parked off to the left to analyze the output beam into x and y -components. The amplitudes in its x and y -channels will be a sum over x' and y' -paths whose amplitudes (1.2.2) are $\langle x'|\Theta_{in}\rangle = \cos(\Theta_{in}-\Theta)$ and $\langle y'|\Theta_{in}\rangle = \sin(\Theta_{in}-\Theta)$ where relative angle is $\Theta_{in}-\Theta$ of Θ_{in} to Θ -analyzer axes- (x',y') related to lab axes by $\langle x|x'\rangle = \cos\Theta = \langle y|y'\rangle$ and $\langle y|x'\rangle = \sin\Theta = -\langle x|y'\rangle$.
 Output x -component is: $\langle x|\Theta_{out}\rangle = \langle x|x'\rangle\langle x'|\Theta_{in}\rangle + \langle x|y'\rangle\langle y'|\Theta_{in}\rangle = \cos\Theta\cos(\Theta_{in}-\Theta) - \sin\Theta\sin(\Theta_{in}-\Theta) = \cos\Theta_{in}$
 Output y -component is: $\langle y|\Theta_{out}\rangle = \langle y|x'\rangle\langle x'|\Theta_{in}\rangle + \langle y|y'\rangle\langle y'|\Theta_{in}\rangle = \sin\Theta\cos(\Theta_{in}-\Theta) + \cos\Theta\sin(\Theta_{in}-\Theta) = \sin\Theta_{in}$.
 (Recall $\cos(a+b) = \cos a \cos b - \sin a \sin b$ and $\sin(a+b) = \sin a \cos b + \cos a \sin b$)
 If the Θ -analyzer vanished the xy -analyzer sees output $\langle x|\Theta_{in}\rangle = \cos\Theta_{in}$ and $\langle y|\Theta_{in}\rangle = \sin\Theta_{in}$ straightaway.
 So $\Theta_{out} = \Theta_{in}$ is unchanged by a "Do-nothing" and having it is the same as having no analyzer.

(1) Optical analyzers in sorter-counter configuration

An analyzer can easily be reduced to a simple sorter-counter of the type discussed in Section 1.2. You just block the ends of the x -high road and the y -low road with counters as shown in Fig. 1.3.3.

$$x\text{-counts} \sim |\langle x|x'\rangle|^2 = \cos^2 \theta = 0.75$$

$$y\text{-counts} \sim |\langle y|x'\rangle|^2 = \sin^2 \theta = 0.25$$

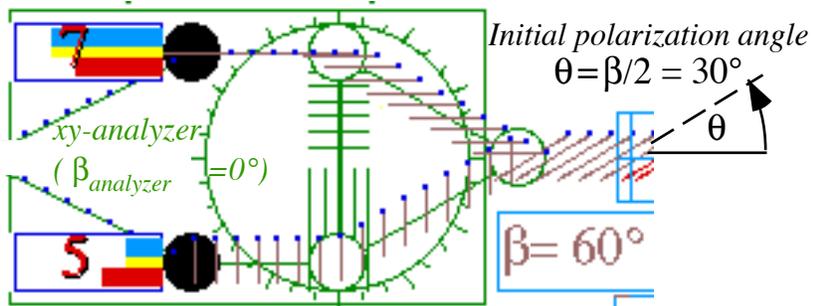


Fig. 1.3.3 Computer sketch of simulated polarization analyzer set up as a sorter-counter

This is the most extreme photon "perturbation" that analyzers can do; it kills any and all photons that venture into this experiment. Not nice! But, one gets an accurate "body-count" to compare with the quantum predictions from (1.2.12) which are shown on the left hand side of Fig. 1.3.3.

The predictions for $(\theta = \beta/2 = 30^\circ)$ are 75% probability for x -polarized photons and 25% probability for y -polarized photons. This particular experiment involving 12 photons came out with 7 counts of x -photons and 5 counts of y -photons. Was that pretty good? Sorry! *It doesn't mean a damn thing!* It could have just as well come out with twelve x -counts and zero y -counts or even *vice-versa*. Or worse, it could have come out 9 to 3 exactly as predicted, and some people who were really stupid would brag that they were doing fabulously great physics. (Then you could ask them what they predicted for 13 photons!)

Quantum predictions of this sort only begin to become meaningful for large numbers of counts. Statistics with small samples is generally not very useful for confirming or disproving a particular theory. Fortunately, there is an endless supply of photons just dying to be sacrificed for your research project.

(2) Optical analyzers in a filter configuration (Polaroid® sunglasses)

Another setting for an analyzer closes one of the paths and leaves the other open. The closed path may have a photon counter as shown below in Fig. 1.3.4. That option does not affect the experiment.

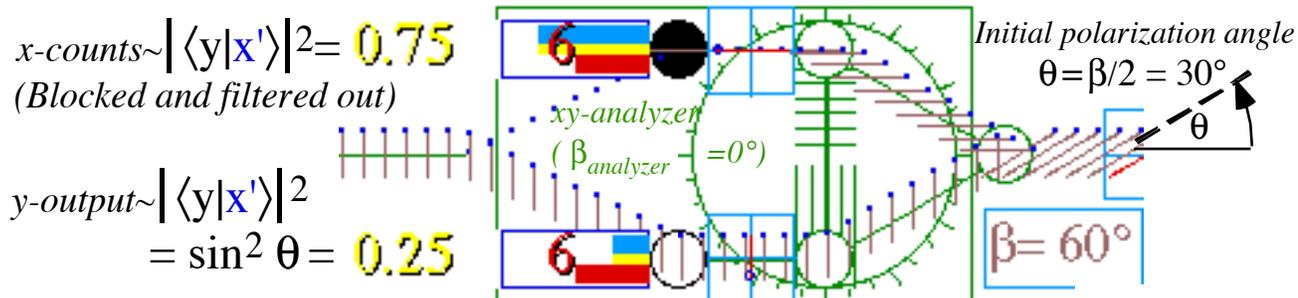


Fig. 1.3.4 Simulated polarization analyzer set up to filter out the x-polarized photons

This experiment simply kills (and counts) all the photons that choose the x-polarized path but lets the photons that choose the y-polarized path go on through. (The latter are counted by default since there are only two choices here.) Our theory says that only 25% should take the y-path on the average just as it did when both paths were blocked in Fig. 1.3.3. This makes sense according to classical arguments since the y-component of the initial 30° polarization E-vector is half of $|E|$ ($\sin 30^\circ = 1/2$) and that corresponds to the one-quarter intensity ($(1/2)^2 = 1/4$) getting through this y-pass filter.

However, any given quantum experiment will probably deviate from the prediction. After twelve "throws" this particular experiment paid off 6 y-photons. That is 50% which is twice the house odds. (Wanna' bet even odds on the next twelve, you old riverboat gambler, you!?)

A lot of money has been "won" by an "experiment" like this, but not by betting on single photon events. Edwin Land sold many Polaroid® sunglasses which pass mostly y-polarized light and block the x-polarized light that accounts for most of the glare reflected from roads by the Brewster effect. You should take heed. Any experiment, like this or ones we will study later, might become a multi-million \$ invention if you use your imagination creatively.

(3) Optical analyzers in the "control" configuration: Half or Quarter wave plates

Now let's see an example of an analyzer configuration that does what analyzers are really intended to do. Recall that a "do-nothing" analyzer was ever so carefully adjusted so the two paths were the same to within a tiny fraction of an optical wave oscillation. This was necessary to assure that the polarization of the input beam is reproduced as perfectly as possible in the output beam.

A "do-something" analyzer requires the same care and precision in adjusting the two paths, but the two paths are given a non-zero difference in optical path length or *relative phase* Ω . If this phase difference can be accurately set and controlled over a range of 2π ($-\pi < \Omega < \pi$) then the analyzer is capable of completely controlling the output polarization. If the x-path gains a half-wave or phase of $\Omega = \pi$ relative to the y-path, then input 30° polarization becomes tilted by an angle of $\theta = 150^\circ$ as shown in Fig. 1.3.5a. Such a device is called a

half-wave plate. If the x -path gains a quarter-wave or phase of $\Omega=\pi/2$ relative to the y -path, then input 30° polarization becomes *elliptical* polarization as shown in Fig. 1.3.5b. This device is called a *quarter-wave plate*.

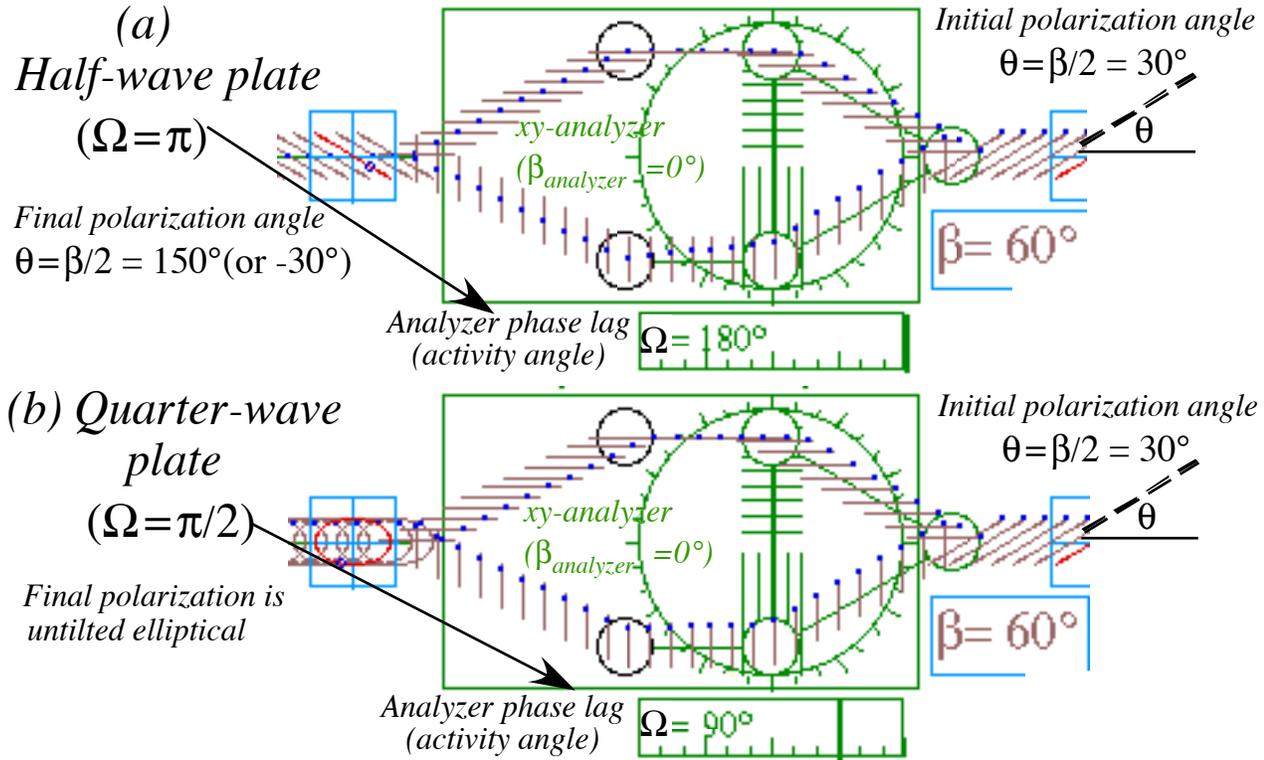


Fig. 1.3.5 Polarization control set to shift phase by (a) Half-wave ($\Omega=\pi$) , (b) Quarter wave ($\Omega=\pi/2$)

To understand what happened in Fig. 1.3.5 consider the input $\theta=30^\circ$ polarization state vector given by the $|x'\rangle$ ket in (1.2.2) or (1.2.4a). Then we shift its x -phase by an angle $\Omega=\pi$ or $\pi/2$, respectively, relative to the y -phase, by multiplying Ψ_x by phase factor $e^{-i\Omega}$.

$$\begin{aligned} \begin{pmatrix} e^{-i\Omega} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} &= \begin{pmatrix} e^{-i\pi} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} & \begin{pmatrix} e^{-i\Omega} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} &= \begin{pmatrix} e^{-i\pi/2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} \\ &= \begin{pmatrix} -\sqrt{3}/2 \\ 1/2 \end{pmatrix} \text{ for } \Omega=\pi, & & = \begin{pmatrix} -i\sqrt{3}/2 \\ 1/2 \end{pmatrix} \text{ for } \Omega=\pi/2. \end{aligned} \quad (1.3.1a) \quad (1.3.1b)$$

Phase shifting is done using 2×2 complex *matrix operators* such as T or the *phase-balanced form* R below.

$$T = \begin{pmatrix} e^{-i\Omega} & 0 \\ 0 & 1 \end{pmatrix}, \text{ or: } R = \begin{pmatrix} e^{-i\Omega/2} & 0 \\ 0 & e^{+i\Omega/2} \end{pmatrix} = e^{+i\Omega/2} T. \quad (1.3.1c)$$

Whenever possible, one tries to represent analyzers by complex operators \mathbf{T} which act on input state kets $|\Psi_{IN}\rangle$ to give the resulting output state kets $|\Psi_{OUT}\rangle$. Phase-balanced forms are designed to preserve overall phase.

$$|\Psi_{OUT}\rangle = \mathbf{T}|\Psi_{IN}\rangle \quad (1.3.2)$$

A classical picture of the resulting time behavior is shown in Fig. 1.3.6. This is obtained by plotting the real parts of the vectors in (1.3.1a-b) after they have been multiplied by the Planck time-frequency factor $e^{-i\omega t}$ from (1.2.7) as follows. The simulations in Fig. 1.3.5 apply the phase-balanced forms R in each case.

$$\begin{pmatrix} \Psi_x(t) \\ \Psi_y(t) \end{pmatrix} = \begin{pmatrix} \Psi_x(0) \\ \Psi_y(0) \end{pmatrix} e^{-i\omega t} = \begin{pmatrix} -\sqrt{3}/2 \\ 1/2 \end{pmatrix} e^{-i\omega t},$$

$$\begin{pmatrix} \text{Re } \Psi_x(t) \\ \text{Re } \Psi_y(t) \end{pmatrix} = \begin{pmatrix} -(\sqrt{3}/2)\cos \omega t \\ (1/2)\cos \omega t \end{pmatrix} \text{ for } \Omega = \pi, \tag{1.3.3a}$$

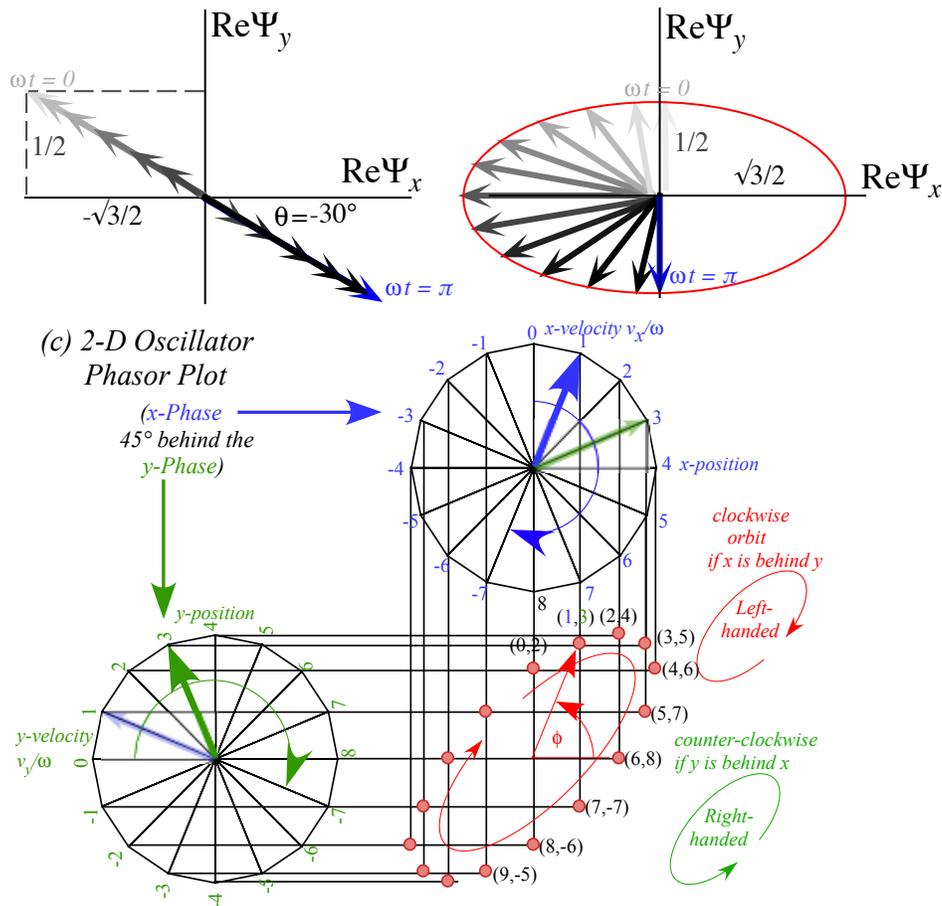
$$\begin{pmatrix} \Psi_x(t) \\ \Psi_y(t) \end{pmatrix} = \begin{pmatrix} \Psi_x(0) \\ \Psi_y(0) \end{pmatrix} e^{-i\omega t} = \begin{pmatrix} -i\sqrt{3}/2 \\ 1/2 \end{pmatrix} e^{-i\omega t}$$

$$\begin{pmatrix} \text{Re } \Psi_x(t) \\ \text{Re } \Psi_y(t) \end{pmatrix} = \begin{pmatrix} (-\sqrt{3}/2)\sin \omega t \\ (1/2)\cos \omega t \end{pmatrix} \text{ for } \Omega = \pi/2. \tag{1.3.3b}$$

The phase factor of (-1) in the x -component of eq. (1.3.3a) simply reflects the polarization oscillation plane through the x -axis as shown in Fig. 1.3.6a and on the left of Fig. 1.3.5a. The phase factor of (i) in the x -component of eq. (1.3.3b) causes the real polarization vector to trace an elliptical path in the xy -plane as shown in Fig. 1.3.6b and on the left (output) of Fig. 1.3.5b.

Note the polar form of i is $e^{i\pi/2}$ with polar angle $\pi/2$ or 90° . Therefore, we say that the oscillation with the i -factor or $\pi/2$ phase is 90° counter-clockwise to an oscillation that has no extra phase factor. An i -factor x -oscillation is 90° behind y but ($-i$) makes x go 90° ahead (giving anti-clockwise rotation as in Fig. 1.3.6b) since the $e^{-i\omega t}$ phase of Planck advances clockwise with time. The factor $e^{-i\Omega}$ with $\Omega=-45^\circ$ gives a tilted clockwise rotating elliptical polarization as in Fig. 1.3.6(c).

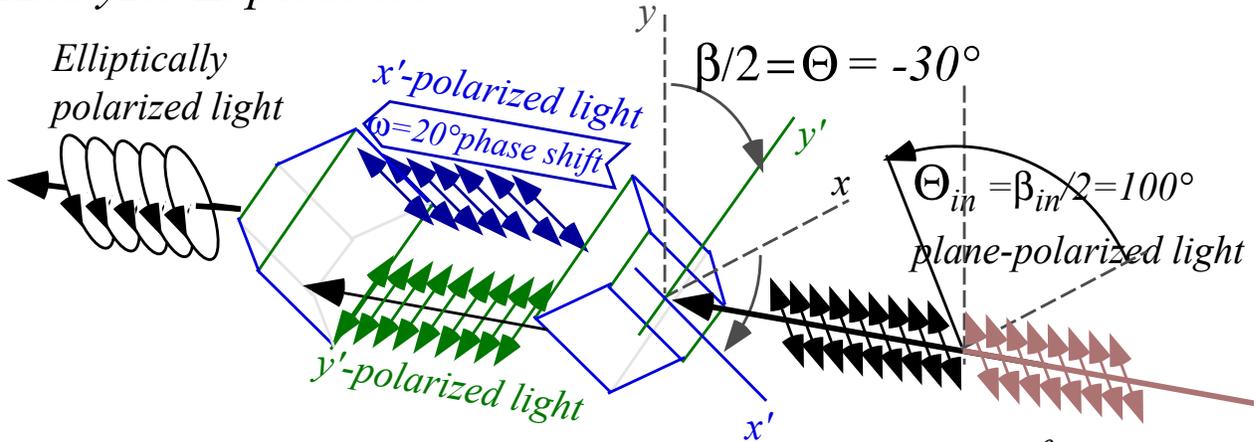
Fig. 1.3.6 Polarization states for (a) Half-wave ($\Omega=\pi$) , (b) Quarter wave ($\Omega=\pi/2$) (c) ($\Omega=-\pi/4$)



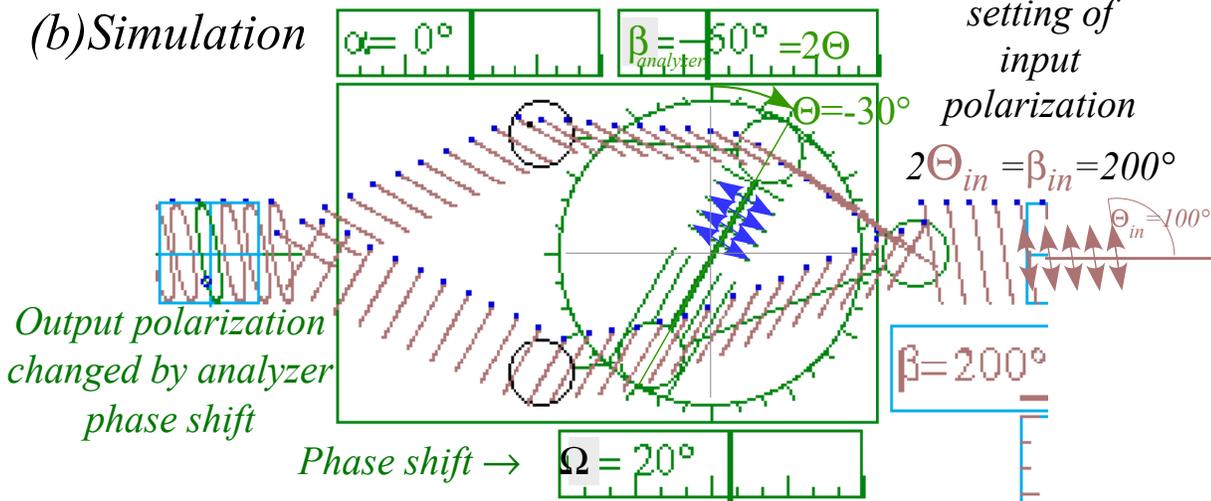
Simulation of Active or "Do-Something" Analyzers

Suppose an active analyzer, that is a "do-something" analyzer, shifts its high beam by a phase angle ($\Omega=20^\circ$) as indicated above the analyzer in the figure below. As a result the initial plane- 100° -polarization coming in gets transformed to an elliptical polarization going out. How do we calculate the Ψ_{out} state now? The key is to put in the Ω -phase shift in the right place.

(a) Analyzer Experiment



(b) Simulation



The output calculation is almost the same as it is for the "do-nothing" analyzer except now an extra phase factor $e^{-i\Omega} = 0.94-i 0.34$ is tacked onto factors for the x' -path as was discussed after (1.3.1).

$$x\text{-output: } \langle x | \Psi_{out} \rangle = \langle x | x' \rangle e^{-i\Omega} \langle x' | \Psi_{in} \rangle + \langle x | y' \rangle \langle y' | \Psi_{in} \rangle = e^{-i\Omega} \cos \Theta \cos(\Theta_{in} - \Theta) - \sin \Theta \sin(\Theta_{in} - \Theta)$$

$$y\text{-output: } \langle y | \Psi_{out} \rangle = \langle y | x' \rangle e^{-i\Omega} \langle x' | \Psi_{in} \rangle + \langle y | y' \rangle \langle y' | \Psi_{in} \rangle = e^{-i\Omega} \sin \Theta \cos(\Theta_{in} - \Theta) + \cos \Theta \sin(\Theta_{in} - \Theta)$$

The numerical results for these amplitudes are as follows. Both of these need to be given in polar form.

$$x\text{-output: } \langle x | \Psi_{out} \rangle = (0.94-i0.34)0.87(-0.64) - (-0.5)(0.77) = -0.140 + i0.189 = 0.235e^{i2.2}$$

$$y\text{-output: } \langle y | \Psi_{out} \rangle = (0.94-i0.34)(-0.5)(-0.64) + (0.87)(0.77) = 0.966 - i0.109 = 0.972e^{-i0.11}$$

The x and y -probabilities are $(0.235)^2=0.055$ (5.5%) and $(0.972)^2=0.945$ (94.5%), respectively. Multiplying the amplitudes by the Planck time phasor $e^{-iEt/\hbar}$ gives their time dependence. (See (1.3.3)) The x -amplitude is running $2.2-(-0.1)=2.3$ radians or 131° behind the y -amplitude. This makes a slightly tilted ellipse. The detailed geometry of polarization ellipsometry is discussed in Chapter 10.

(4) Optical analyzers in a "peeking" configuration

The following is a preliminary discussion about one of the most peculiar aspects of quantum theory. It falls under a large and fairly nebulous topic of *quantum measurement* or *quantum observation*. The main idea, which we will discuss later on, is that the information about a given system comes in quantum chunks just like its energy and momenta ; indeed, information is energy and momenta since information requires frequency bandwidth. So information about a quantum system comes at a price, and the price is that the energy and state of the system is affected at the quantum level.

Suppose a new analyzer is constructed with both ports open just like the original "do-nothing" analyzer first shown in Fig. 1.3.1 and 1.3.2. Only the new analyzer has a "spy" or an "x-eye" which "peeks" and looks out for a passing *x*-photon as shown in Fig. 1.3.5. It may also have another "y-eye" which "peeks" at the lower *y*-path to check if a *y*-photon went by. However, for a two-state system, one good eye is assumed to be enough to tell which way the photon went.

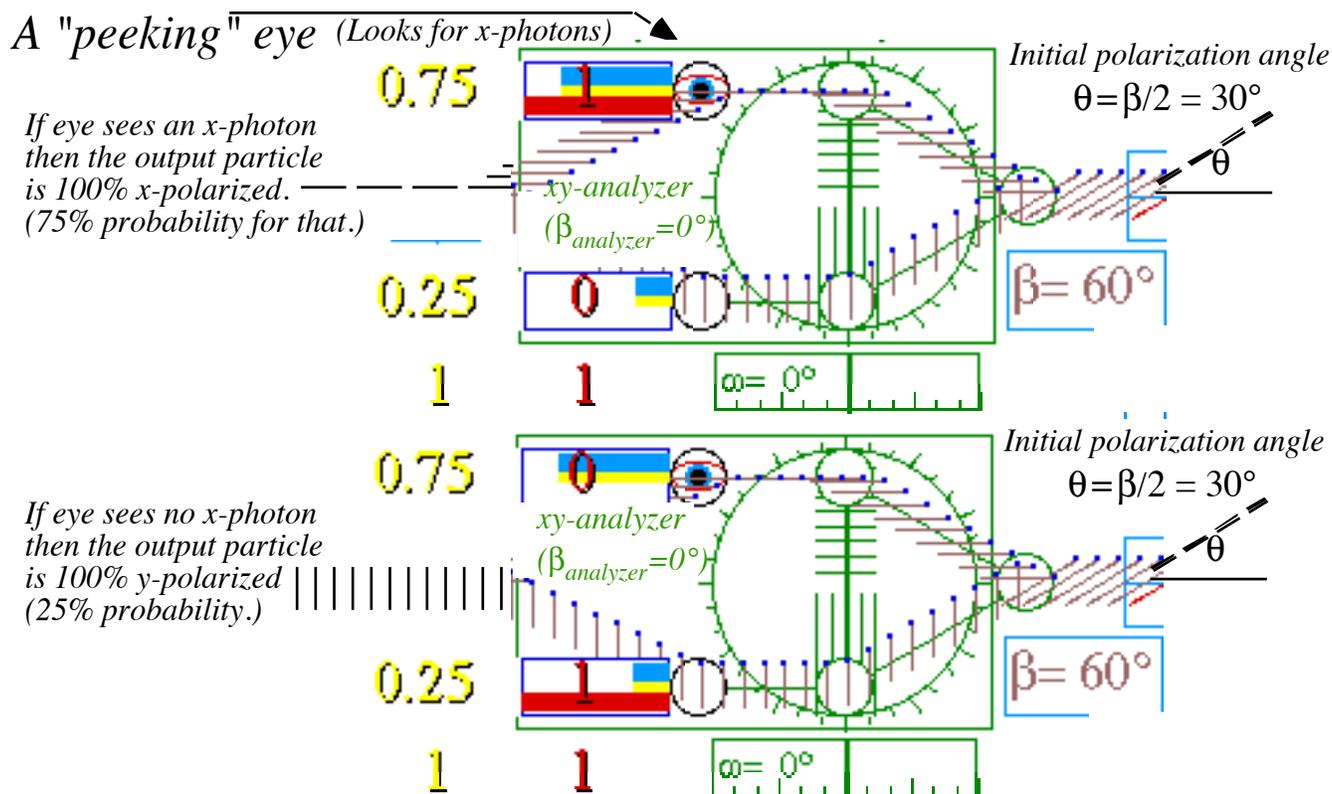


Fig. 1.3.7 Simulated polarization analyzer set up to "peek" if the photon is *x*-or *y*-polarized

It should be emphasized that the "peeking" is as delicate as possible, a sort of "ideal peeking." It does not change the *x*-polarization of any of the *x*-photons it "sees." Nor does it change the *y*-polarization of any of the *y*-photons. Neither does it alter the predictions that our 30° polarized input beam would, if so analyzed, yield 75% of the photons to be *x*-polarized and the remaining 25% to be *y*-polarized.

However, it cannot avoid altering the state of the incoming beam. If a *x*-photon is "seen" then that is exactly what comes out; the output photon is 100% *x*-polarized. If no *x*-photon is seen then the output is a 100% *y*-polarized photon. This analyzer can only output *x*-polarized or *y*-polarized photons, and nothing in

between. Furthermore, after counting many photons the x -polarized photons will, on the average, be 75% of the population and the y -photons will account for the remaining 25%.

Such "perfect-peeking" is like a sorter counter in Fig. 1.3.3 which, after each counter records a count, recreates a (randomly phased) photon with the same polarization as the one it absorbed.

This output is very different from that of the "do-nothing" analyzer whose output was entirely made up of 30° polarized photons, that is each photon had 75% intensity in the x -direction ($\Psi_x = \sqrt{3}/2$, $|\Psi_x|^2 = 0.75$) and 25% intensity in the y -direction ($\Psi_y = 1/2$, $|\Psi_y|^2 = 0.25$). In contrast, the "peeked" photons are either all x and no y ($|\Psi_x|^2 = 0.75$, and $|\Psi_y|^2 = 0.00$) as are 75% of the photons, or else all y and no x ($|\Psi_x|^2 = 0.00$, and $|\Psi_y|^2 = 0.25$) as are the remaining 25%. It's as though peeking produces two separate beams within the final beam neither of which have 100% probability individually, but together their probabilities add to 1. (This is discussed in greater detail after Fig. 1.3.9b just ahead.)

One of the most bizarre aspects of quantum theory is that by looking to see if a particle chooses the x -path or the y -path we seem to cause it become exactly one of the possible results that we are looking for; in this case, either x or else y but not a combination of both. This is sometimes called *wave amplitude collapse* since the amplitude for the "other" path (or paths for more than two states) seems to instantly be shut off. However, this little bit of jargon is highly suspicious and perhaps it should be deleted from our vocabulary.

It may seem paradoxical that by turning on the " x -eye" and seeing an x -photon on the x -path instantly shuts off any chance that the " y -eye" will see a photon down on the y -path at that moment or until another photon comes along. Could the x -path observer use amplitude collapse to communicate instantaneously (or faster than light) with the distant y -path observer?

No! Remember the x -photon that the " x -eye" recorded was not a sure thing until it actually happened, rather it had a probability of only 75%. The y -path observer is not going to notice any "shut-off" if that particular instant doesn't give him a y -photon count. The poor y -observer is only getting 25% of the photons, anyway. Furthermore, there is a 25% probability that after the x -observer turns on his " x -eye" the first photon goes instead to the y -observer. What kind of message is that? It is a garbled one with exactly zero information. Amplitude collapse is an example of what has been called "spooky action at a distance." So far "spooky action" has not been made to work for us in spite of the fact that modern quantum theory has many examples of this sort of thing for over a century.

(b) Effects of "peeking" : Coherent versus incoherent beams

Understanding the effects of "peeking" is one of the most difficult parts of fundamental quantum mechanics. To help in this we now compare the output of an xy -*"Do nothing"* analyzer like the one in Fig. 1.3.2 with that of an xy -*"Peeker"* analyzer like the one in Fig. 1.3.7. A second "counting" analyzer tilted by an angle of $\beta/2 = 30^\circ$ is set to count the output of the "Do-nothing" in the upper left half of Fig. 1.3.8 and the same $\beta/2 = 30^\circ$ counter counts output of the xy -*"Peeker"* in the lower left half. For each case the initial beam (Extreme right of Fig. 1.3.8) is the same $\beta/2 = 30^\circ$ input that has been used in all the previous cases, but the resulting counts on the extreme left are very different. It is important to see why.

The $\beta/2 = 30^\circ$ tilted counter will sort any beam into $\beta/2 = 30^\circ$ tilted xy -polarized beams with a $|x'\rangle$ -beam going to the upper counter and a $|y'\rangle$ -beam going to the lower counter. The initial input beam on the extreme right of Fig. 1.3.8 is a pure $|x'\rangle$ -beam, and the "Do-nothing" analyzer, true to its name, does nothing to change $|x'\rangle$ -polarization. So all photons go straight to the upper path and into the $|x'\rangle$ counter 100% of the time. The $|y'\rangle$ -counter gets exactly zero counts.

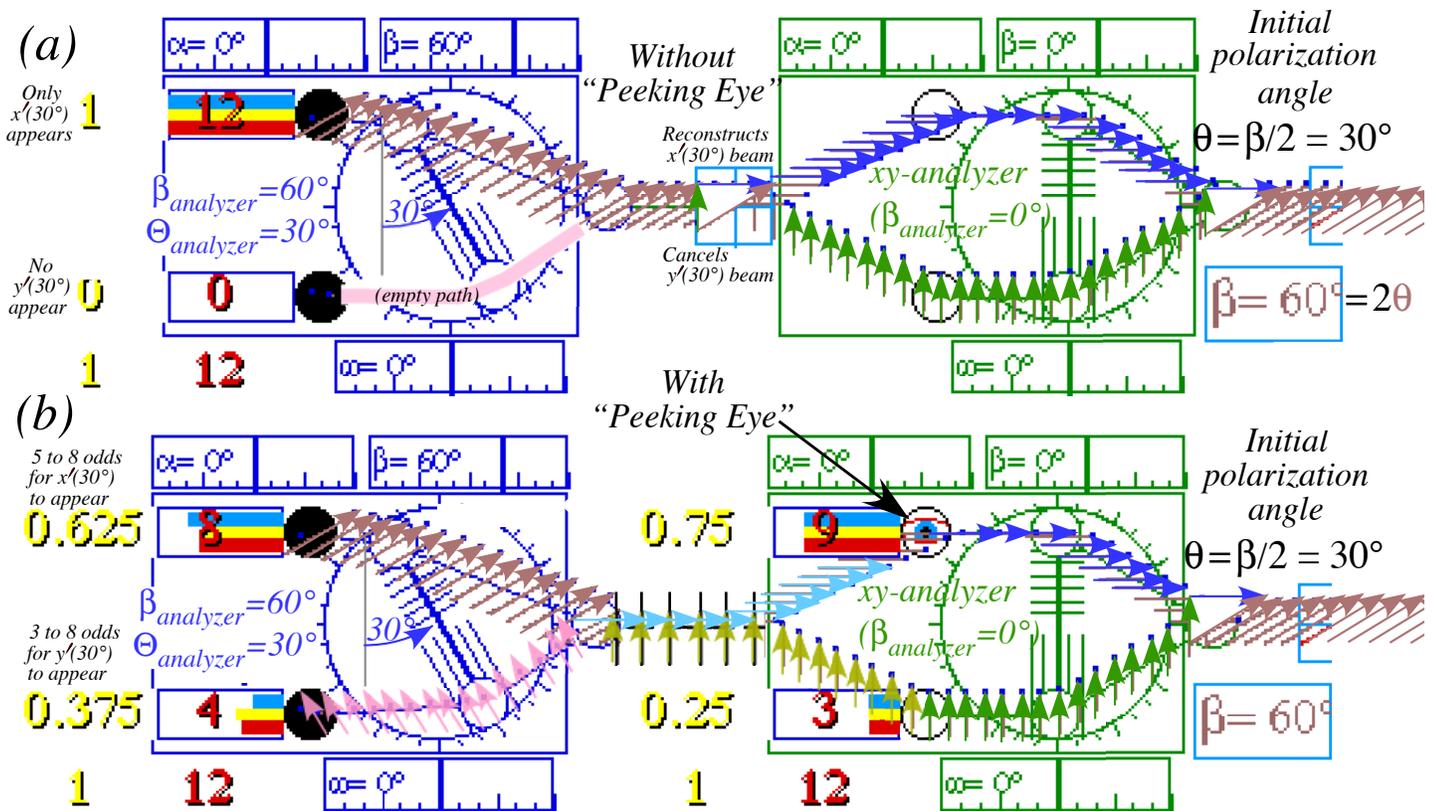


Fig. 1.3.8 Output with $\beta/2=30^\circ$ input to: (a) Coherent xy -*"Do nothing"* or (b) Incoherent xy -*"Peeking"* devices

However, the 'Peeking' analyzer in the lower right half of Fig. 1.3.8 will put out two kinds of polarized beams; an x -polarized beam with 75% of the intensity and a y -polarized beam with 25% of the intensity, as we discussed earlier. (Sec. 1.3 a(5)). These each get sorted and counted separately.

A sketch in Fig. 1.3.9 below of the polarization vector components or amplitudes can explain this. The upper right hand portion shows the initial $|x'\rangle$ -beam polarization being resolved into its x - and y -components

so they split into the $|x\rangle$ -beam and the $|y\rangle$ -beam which are then recombined to make a final $|x'\rangle$ -beam output on the upper left hand side. The orthogonal $|y'\rangle$ component has zero amplitude.

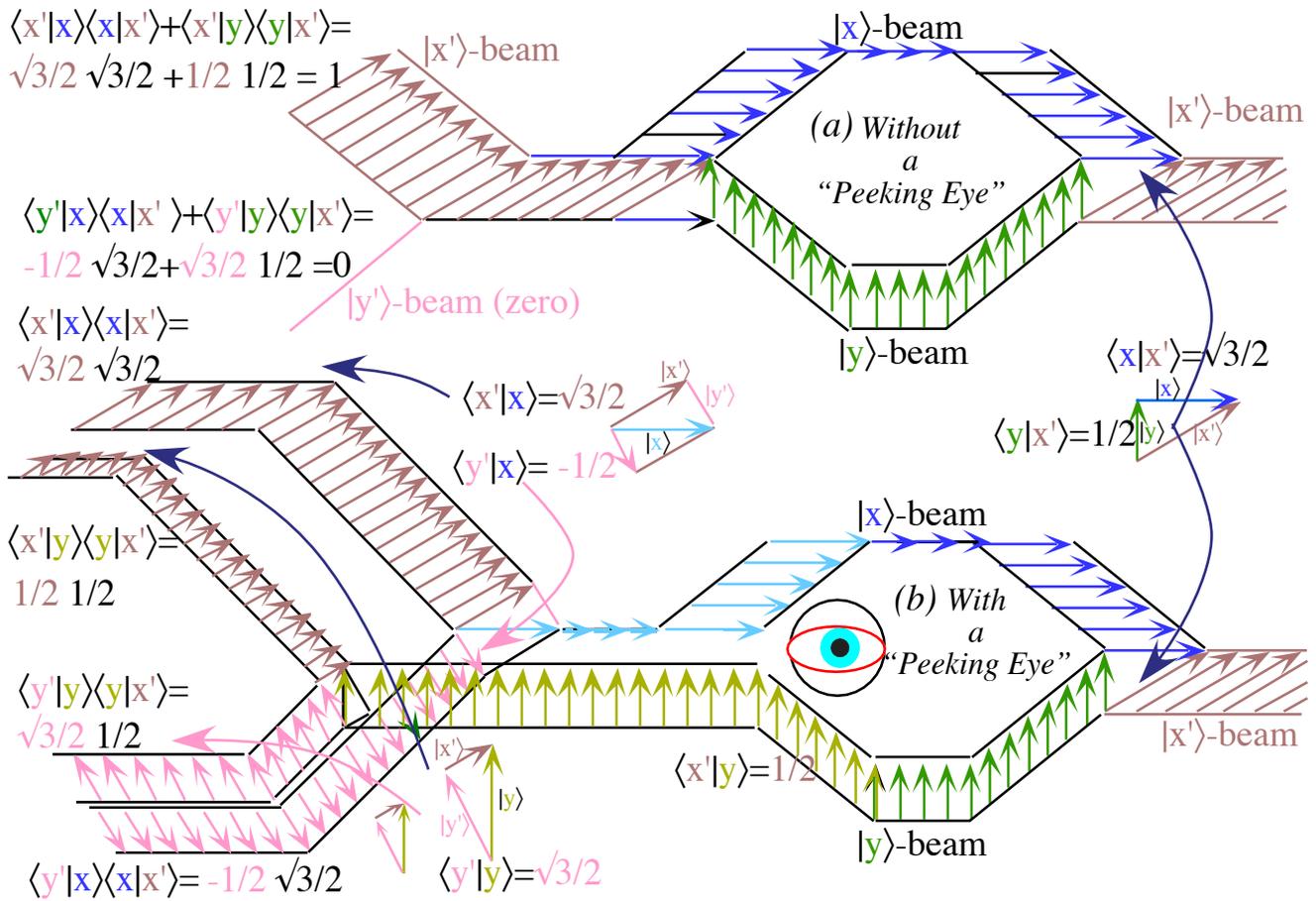


Fig. 1.3.9 Beams-amplitudes of (a) xy -“Do nothing” and (b) xy -“Peeking” analyzer each with $|x\rangle$ input

A more detailed view of components and amplitudes is shown in the lower portion where we imagine that the "peeking eye" has caused the two beams to be distinguished. In other words, something has happened to make the $|x\rangle$ -beam or $|y\rangle$ -beams "dirty" so that they cannot recombine to make a pure $|x'\rangle$ -beam going up to the $|x'\rangle$ counter and a zero beam going down to the $|y'\rangle$ counter. This "dirt" is simply a *random phase* which an "eye" must add to every photon it "sees." Without this "dirt" the lower two amplitudes cancel to 0 through perfect *destructive interference* while the upper beams add up to 1 through perfect *constructive interference*. ("Constructive interference" is an oxymoron in the "dirty" classical world we're used to, but it is absolutely essential for understanding wave behavior in the quantum world.)

(1) Amplitude products

We now consider the effect of successive analyzers such as the example in Fig. 1.3.8 and 9 where an xy -analyzer acts on an initial state $|x\rangle$ and is followed by an $x'y'$ -counter. This example provides the opportunity to introduce more of the fundamentals of quantum analysis and mathematics.

According to our discussion in Section 1.2 b, when a beam in state $|\Psi\rangle$ enters a sorter it gets sorted into a series of beams $\{|b_1\rangle, |b_2\rangle, \dots\}$ of amplitude $\langle b_1|\Psi\rangle, \langle b_2|\Psi\rangle, \dots$, respectively. (So far only two-state systems have been shown.) If one of these beams (say, $|b_m\rangle$) encounters another analyzer or sorter it gets sorted into another series of beams $\{|c_1\rangle, |c_2\rangle, \dots\}$ of amplitude $\langle c_1|b_m\rangle\langle b_m|\Psi\rangle, \langle c_2|b_m\rangle\langle b_m|\Psi\rangle, \dots$, respectively. Note: each sorter amplitude $\langle c_1|b_m\rangle, \langle c_2|b_m\rangle, \dots$, gets multiplied by the input amplitude $\langle b_m|\Psi\rangle$. This is because polarization response is *linear* in the input amplitude, as indeed, are all analogous quantum processes. Output is proportional to input.

This process happens again each time a beam runs into another sorter and makes another set of "baby beams." For example, an amplitude of the form shown in Fig. 1.3.10

$$\langle e_p | D_o C_n B_m | \Psi \rangle = \langle e_p | d_o \rangle \langle d_o | c_n \rangle \langle c_n | b_m \rangle \langle b_m | \Psi \rangle \quad (1.3.4)$$

is the contribution to the $|e_p\rangle$ beam of the e-sorter by a particle that went through $|d_o\rangle$ beam of a d-sorter after passing the $|c_n\rangle$ beam of the c-sorter, after passing the $|b_m\rangle$ beam of the b-sorter from the input $|\Psi\rangle$ beam. We just read this series backwards from finish to start. Recall that you read amplitudes like Hebrew, right to left, in going forward from start to finish or (correct terminology) initial to final.

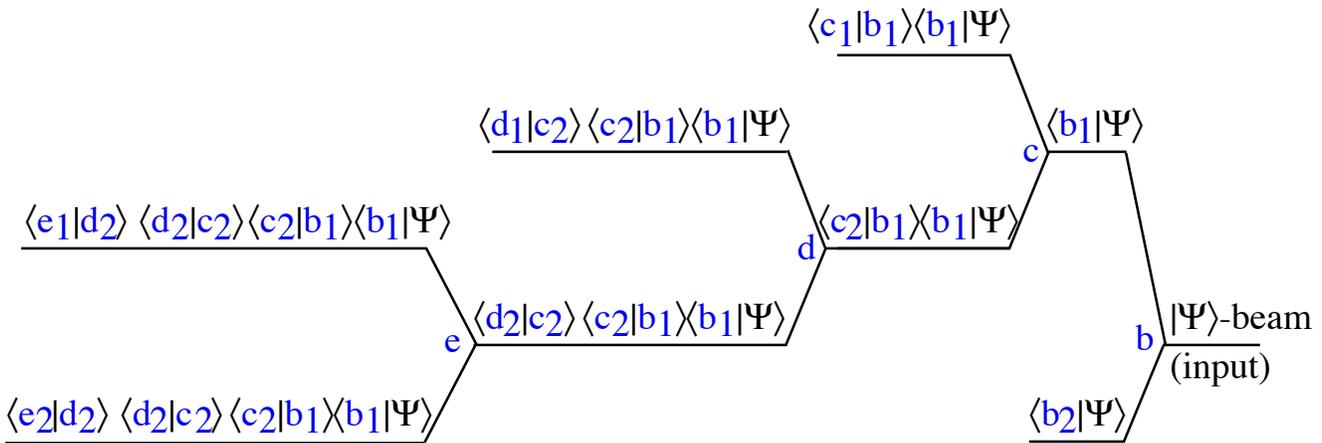


Fig. 1.3.10 Beams-amplitude products for successive beam sorting

(2) Amplitude sums

Analyzers recombine their sorted beams and the resulting analyzer output amplitude is a sum of the amplitudes of its beams. For example, the final amplitude of the x' or y' -counters in the upper left hand part of Fig. 1.3.9 is given by the following sum. (Recall that the initial state is $|\Psi\rangle = |x'\rangle$.)

$$\begin{aligned} \text{Amp. at } x'\text{-counter} &= \langle x'|x\rangle\langle x|\Psi\rangle + \langle x'|y\rangle\langle y|\Psi\rangle = \langle x'|x\rangle\langle x|x'\rangle + \langle x'|y\rangle\langle y|x'\rangle \\ \text{Amp. at } y'\text{-counter} &= \langle y'|x\rangle\langle x|\Psi\rangle + \langle y'|y\rangle\langle y|\Psi\rangle = \langle y'|x\rangle\langle x|x'\rangle + \langle y'|y\rangle\langle y|x'\rangle \end{aligned} \quad (1.3.5)$$

Amplitudes are diagrammed as 30° - 60° right triangle segments in Fig. 1.3.9 according to Sec. 1.2a

$$\begin{aligned} \begin{pmatrix} \langle x'|x\rangle & \langle x'|y\rangle \\ \langle y'|x\rangle & \langle y'|y\rangle \end{pmatrix} &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, & \begin{pmatrix} \langle x|x'\rangle & \langle x|y'\rangle \\ \langle y|x'\rangle & \langle y|y'\rangle \end{pmatrix} &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{pmatrix}, \text{ for } \theta=30^\circ, & = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}, \text{ for } \theta=30^\circ \end{aligned} \quad (1.3.6)$$

(Notice that the two transformation matrices are inverses of each other.) Substituting in (1.3.5) gives the correct amplitudes for the "do nothing" analyzer.

$$\text{Amp. at } x'\text{-counter} = \langle x'|x\rangle\langle x|x'\rangle + \langle x'|y\rangle\langle y|x'\rangle = \frac{3}{4} + \frac{1}{4} = 1 \quad (1.3.7a)$$

$$\text{Amp. at } y'\text{-counter} = \langle y'|x\rangle\langle x|x'\rangle + \langle y'|y\rangle\langle y|x'\rangle = -\frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{4} = 0 \quad (1.3.7b)$$

The two pairs of terms of (1.3.7) are sketched by two pairs of parallel beams in the lower left of Fig. 1.3.9. For an ideal "do nothing" analyzer, the first two in (1.3.7a) sum to 1 to give the same x' -beam that came as input, but the second two in (1.3.7b) cancel causing the y' -beam to vanish.

(3) Random phase effects ("Dirty" beams)

However, if any of the four beams is "dirty", that is, has a random phase then the counter output is very different. Suppose the " x -eye" tags each x -photon with a phase $e^{i\phi}$ where ϕ is a different random number for each photon so ϕ ranges over the unit circle ($-\pi < \phi < \pi$). Then (1.3.7) becomes

$$\text{Amp. at } x'\text{-counter} = \langle x'|x\rangle\left(e^{i\phi}\right)\langle x|x'\rangle + \langle x'|y\rangle\langle y|x'\rangle = \frac{3}{4}\left(e^{i\phi}\right) + \frac{1}{4} \quad (1.3.8a)$$

$$\text{Amp. at } y'\text{-counter} = \langle y'|x\rangle\left(e^{i\phi}\right)\langle x|x'\rangle + \langle y'|y\rangle\langle y|x'\rangle = -\frac{\sqrt{3}}{4}\left(e^{i\phi}\right) + \frac{\sqrt{3}}{4} \quad (1.3.8b)$$

Recall that statistical probabilities and count rates are determined by the absolute square of amplitudes according to (1.2.12). The count probabilities from (1.3.8) are as follows.

$$\begin{aligned} x'\text{-count probability} &= \left| \frac{3}{4}\left(e^{i\phi}\right) + \frac{1}{4} \right|^2 = \left(\frac{3}{4}\left(e^{i\phi}\right) + \frac{1}{4} \right)^* \left(\frac{3}{4}\left(e^{i\phi}\right) + \frac{1}{4} \right) \\ &= \frac{3^2}{4^2} + \frac{1^2}{4^2} + \frac{3}{4^2}\left(e^{-i\phi} + e^{i\phi}\right) = \frac{5}{8} + \frac{3}{16}\left(e^{-i\phi} + e^{i\phi}\right) \end{aligned} \quad (1.3.9a)$$

$$\begin{aligned} y'\text{-count probability} &= \left| -\frac{\sqrt{3}}{4}\left(e^{i\phi}\right) + \frac{\sqrt{3}}{4} \right|^2 = \left(-\frac{\sqrt{3}}{4}\left(e^{i\phi}\right) + \frac{\sqrt{3}}{4} \right)^* \left(-\frac{\sqrt{3}}{4}\left(e^{i\phi}\right) + \frac{\sqrt{3}}{4} \right) \\ &= \frac{3}{4^2} + \frac{3}{4^2} - \frac{3}{4^2}\left(e^{-i\phi} + e^{i\phi}\right) = \frac{3}{8} - \frac{3}{16}\left(e^{-i\phi} + e^{i\phi}\right) \end{aligned} \quad (1.3.9b)$$

Recall also, that probabilities are useful predictors only for large numbers of trials, that is, after many photons, so the average of $e^{i\phi}$ or $e^{-i\phi}$ approaches zero if ϕ is random. Then (1.3.8) reduces to

$$x'\text{-count probability} = 5/8 = 0.625, \quad y'\text{-count probability} = 3/8 = 0.375, \quad (1.3.9c)$$

which agrees with the simulation in the lower half of Fig. 1.3.8. Main idea: "Dirty" beams don't interfere.

(4) Summing amplitudes or probabilities?

This example exposes one of the important differences between quantum theory and classical physics. Let's assume that the act of "peeking" or "measurement" will, at the very least, jiggle the phase of the object of the measurement. Let's see how "Jiggling" causes count rates to exactly match what classical probability analysis would give. According to classical reasoning the probability that a x'-photon shows up in the x'-counter via an xy-analyzer equals a sum of probability products $P(x' to x)$ times $P(x to x')$ for the x-path and $P(x' to y)$ times $P(y to x')$ for the y-path. Read products below from right to left. (QM uses Hebrew style.)

$$\left(\begin{array}{c} \text{Probability that} \\ \text{photon in x'-input} \\ \text{becomes} \\ \text{photon in x'-counter} \end{array} \right)_{\text{classical}} = \left(\begin{array}{c} \text{probability that} \\ \text{photon in x-beam} \\ \text{becomes} \\ \text{photon in x'-counter} \end{array} \right) * \left(\begin{array}{c} \text{probability that} \\ \text{photon in x'-input} \\ \text{becomes} \\ \text{photon in x-beam} \end{array} \right) + \left(\begin{array}{c} \text{probability that} \\ \text{photon in y-beam} \\ \text{becomes} \\ \text{photon in x'-counter} \end{array} \right) * \left(\begin{array}{c} \text{probability that} \\ \text{photon in x'-input} \\ \text{becomes} \\ \text{photon in y-beam} \end{array} \right)$$

Now replace the probability factors by their squared-amplitude values $P(x to x') = |\langle x'|x \rangle|^2$, etc..

$$\left(\begin{array}{c} \text{Probability that} \\ \text{photon in x'-input} \\ \text{becomes} \\ \text{photon in x'-counter} \end{array} \right)_{\text{classical}} = \left(|\langle x'|x \rangle|^2 \right) * \left(|\langle x|x' \rangle|^2 \right) + \left(|\langle x'|y \rangle|^2 \right) * \left(|\langle y|x' \rangle|^2 \right) = \left(\left| \frac{\sqrt{3}}{2} \right|^2 \right) * \left(\left| \frac{\sqrt{3}}{2} \right|^2 \right) + \left(\left| \frac{-1}{2} \right|^2 \right) * \left(\left| \frac{1}{2} \right|^2 \right) = \frac{5}{8}$$

This what we got in (1.3.9) by ignoring the phase terms. Indeed, the square of (1.3.8a) is as follows.

$$\left(\begin{array}{c} \text{Quantum probability} \\ \text{at x'-counter} \end{array} \right) = \left| \langle x'|x \rangle (e^{i\phi}) \langle x|x' \rangle + \langle x'|y \rangle \langle y|x' \rangle \right|^2$$

$$= \left| \langle x'|x \rangle \langle x|x' \rangle \right|^2 + \left| \langle x'|y \rangle \langle y|x' \rangle \right|^2 + e^{-i\phi} \langle x'|x \rangle^* \langle x|x' \rangle^* \langle x'|y \rangle \langle y|x' \rangle + e^{i\phi} \langle x'|x \rangle \langle x|x' \rangle \langle x'|y \rangle^* \langle y|x' \rangle^*$$

$$= (\text{classical probability}) + (\text{Phase-sensitive or quantum interference terms})$$

(1.3.10)

The difference between quantum and classical probability predictions comes down to what we call phase-sensitive or *quantum interference* terms. Classical and quantum predictions differ after many counts only if phase factors $e^{i\phi}$ do not average to zero. Otherwise the quantum results reduce to the classical ones.

To summarize: Quantum probability is a square of the sum of probability amplitudes for all indistinguishable paths as in the following example.

$$\left(\begin{array}{c} \text{Quantum probability} \\ \text{at } x'\text{-counter} \end{array} \right) = \left| \langle x'|x\rangle\langle x|x'\rangle + \langle x'|y\rangle\langle y|x'\rangle \right|^2 \tag{1.3.11a}$$

Classical probability has a smaller number of terms from this that are just a sum of the squares of the probability amplitudes for all paths, distinguishable or not, as in the following example.

$$\left(\begin{array}{c} \text{Classical probability} \\ \text{at } x'\text{-counter} \end{array} \right) = \left| \langle x'|x\rangle\langle x|x'\rangle \right|^2 + \left| \langle x'|y\rangle\langle y|x'\rangle \right|^2 \tag{1.3.11b}$$

Here is a general rule : If you know what path a system took, (or better, if Mother Nature knows) don't bother with the full quantum square-of-a-sum. Just take the sum-of-the-squares because it is quite certain that the interference terms will average to zero.

Further comments about complex amplitudes

The polarization amplitudes used so far are idealized in a number of ways. The examples given in Sec. 1.2a are real sines and cosines that depend on geometry (tilt angles) alone. As noted in Sec 1.2b time evolution adds complex factors. Also amplitudes for anomalous polarization response (such as occurs near resonance) are complex because the output is not always in phase with the input. Many interesting quantum processes involve relative phases that are extremely sensitive to input parameters.

(c) Electron polarization analyzers

Electron or ion beam polarization analyzers are definitely high-tech experiments. High vacuums are essential and sophisticated magnetic steering fields are needed to level off and recombine the sorted split beams. A rough sketch of an ideal ion beam analyzer is shown in Fig. 1.3.11 below. It is analogous to the optical polarization analyzer shown before in Fig. 1.3.1.

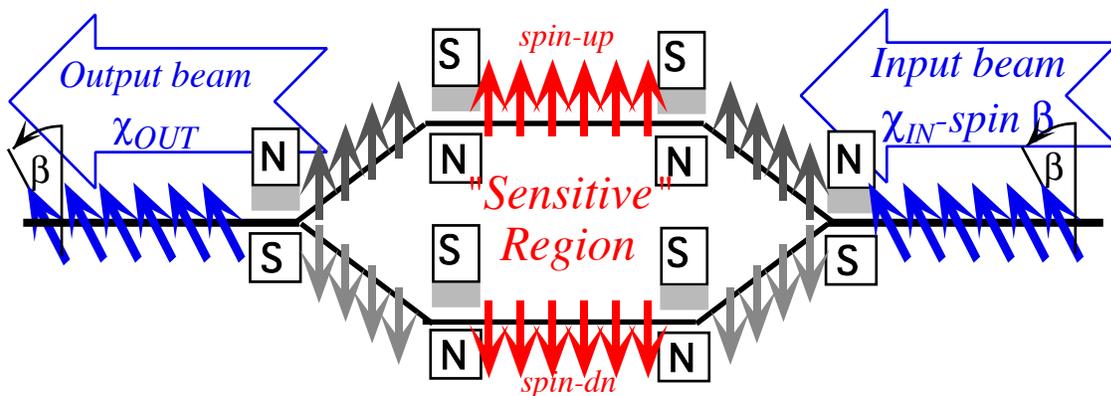


Fig. 1.3.11 Anatomy of ideal electron or ion spin polarization analyzer

In analogy with the optical case, one tests it in a "do-nothing" configuration. It should be possible to adjust it so all input electron spin states come out the same in the output beam, that is, $|\chi_{OUT}\rangle = |\chi_{IN}\rangle$ for all possible $|\chi_{IN}\rangle$. Once this is achieved, then changes can be made in the "sensitive" region where we can "do things" to the separated spin-up and spin-dn beams before they try to recombine.

Appendix 1.A. Review of Complex Algebra

Complex algebra depends on ($i^2=-1$) and the DeMoivre identities (1) or their inverses (2).

$$e^{i\omega t} = \cos \omega t + i \sin \omega t, \quad (1.A.1a)$$

$$\cos \omega t = (e^{i\omega t} + e^{-i\omega t})/2 \quad (1.A.2a)$$

$$e^{-i\omega t} = \cos \omega t - i \sin \omega t, \quad (1.A.1b)$$

$$\sin \omega t = (e^{i\omega t} - e^{-i\omega t})/2i \quad (1.A.2b)$$

This allows any complex number z to be written in *Cartesian* form ($z=x + iy$) or else *polar* form ($z= re^{i\theta}$)

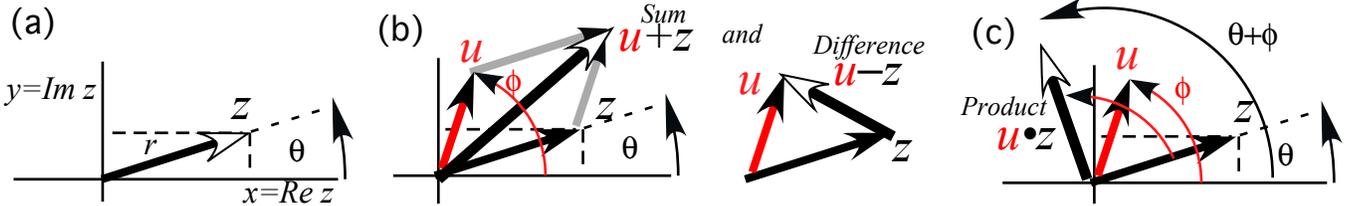
where: $x = \cos \theta \quad (1.A.3a)$

$$r = \sqrt{x^2 + y^2} \quad (1.A.4a)$$

$$y = \sin \theta \quad (1.A.3b)$$

$$\theta = ATAN2(y,x)=\tan^{-1}(y/x) \quad (1.A.4b)$$

are polar-Cartesian coordinate transformations. Caution! Use *ATAN2*, not \tan^{-1} to assure correct angle.



Linear operations (such as addition or subtraction) favor use of Cartesian forms.

$$z+z' = (x+x') + i (y+y')$$

$$(1.A.5a)$$

$$z-z' = (x-x') + i (y-y')$$

$$(1.A.5b)$$

Non-linear operations (such as multiplication or division) favor the use of polar forms.

$$z \cdot z' = (re^{i\theta}) \cdot (r'e^{i\theta'}) = r \cdot r' e^{i(\theta+\theta')} \quad (1.A.6a)$$

$$z/z' = (re^{i\theta})/(r'e^{i\theta'}) = r/r' e^{i(\theta-\theta')} \quad (1.A.6b)$$

Multiplication of $z=re^{i\theta}$ by unitary $u = e^{i\phi}$ (for which $r_u^2 = |u \cdot u| = 1$) gives a rotation of z as seen here.

$$z' = u \cdot z = (e^{i\phi}) \cdot (re^{i\theta}) = re^{i(\theta+\phi)} \quad (1.A.6c)$$

In Cartesian form this equation is a matrix rotation operation.

$$z' = x' + iy' = (\cos \phi + i \sin \phi) \cdot (x + iy) = re^{i(\theta+\phi)} \quad (1.A.7a)$$

$$z' = x' + iy' = (x \cos \phi - y \sin \phi) + i(x \sin \phi + y \cos \phi) = r(\cos(\theta+\phi) + i \sin(\theta+\phi)) \quad (1.A.7b)$$

Equating real parts gives x -rotation and equating imaginary parts gives y -rotation.

$$x' = x \cos \phi - y \sin \phi = r \cos \theta \cos \phi - r \sin \theta \sin \phi = r \cos(\theta+\phi) \quad (1.A.8a)$$

$$y' = x \sin \phi + y \cos \phi = r \cos \theta \sin \phi + r \sin \theta \cos \phi = r \sin(\theta+\phi) \quad (1.A.8b)$$

Angle-sum trig identities for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$ are re-derived in an “automatic” trigonometry.

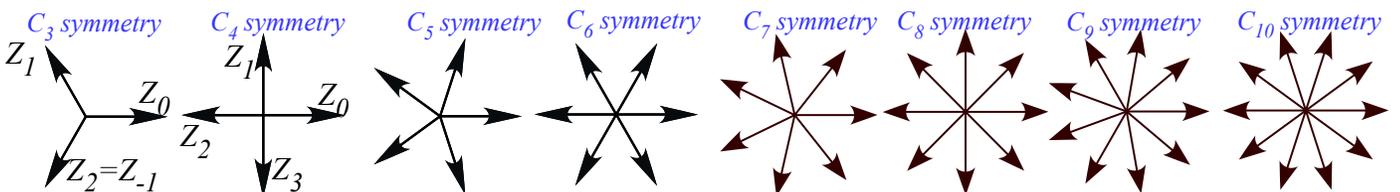
Roots of complex numbers are no problem. For example, solving $z^n = 1$ for $z = (1)^{1/n}$ is simple if we write a polar form of unity: $1 = 1 \cdot e^{2\pi i}$. The zeroth z_0 root is 1 , and the first root z_1 is

$$z_1 = (1)^{1/n} = (e^{2\pi i})^{1/n} = e^{2\pi i/n} \quad (1.A.9)$$

There are n roots $z_k = z_1^k$ because if (z_1^n) equals one, so does $(z_1^n)^2 = (z_1^2)^n$ and $(z_1^3)^n \dots$ and so on.

$$z_1 = e^{2\pi i/n}, z_2 = (z_1)^2 = e^{2(2\pi i/n)}, z_3 = (z_1)^3 = e^{3(2\pi i/n)}, \dots, z_n = (z_1)^n = 1 = z_0 \quad (1.A.10)$$

If these are plotted they form an n -sided regular polygon. Such diagrams help to do z_k arithmetic.



Quantum theory is intimately connected to symmetry, and this arithmetic is the foundation of symmetry analysis and Fourier theory.

Appendix 1.B. Complex Response of Oscillators

To learn quantum theory it helps to know oscillators. A classical stimulated harmonic oscillator equation

$$\frac{d^2 z}{dt^2} + 2\Gamma \frac{dz}{dt} + \omega_0^2 z = a \quad (1.B.1)$$

is solved by complex functions $e^{-i\omega t}$ describing a *monochromatic* (single frequency ω_s) stimulus

$$a(t) = a(0) e^{-i\omega_s t} . \quad (1.B.2)$$

as well as a response at the same frequency if its amplitude is proportional to that of the stimulus.

$$z_{\text{response}}(t) = G_{\omega_0}(\omega_s) a(t) \quad (1.B.3)$$

The complex proportionality factor G depends upon the stimulus frequency ω_s , the natural frequency ω_0 , and damping constant Γ , only. Because the equation is linear and time independent the G factor should not depend on the amplitude A_s of the stimulus. It may help to think of the oscillator as a 'black box' that responds linearly to input as shown below in Fig. 1.B.1.

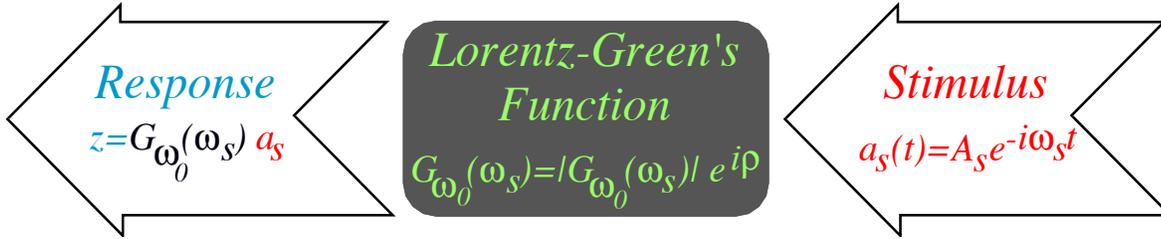


Fig.1.B.1 Black-box diagram of oscillator response to monochromatic stimulus

Now we substitute z_{response} into the classical oscillator equation of motion (1.B.1) and solve for *Lorentzian response function or classical Green's function* $G_{\omega_0}(\omega_s)$ of frequency ω_s .

$$G_{\omega_0}(\omega_s) = \frac{1}{\omega_0^2 - \omega_s^2 - i2\Gamma\omega_s} = \text{Re } G_{\omega_0}(\omega_s) + i \text{Im } G_{\omega_0}(\omega_s) = |G_{\omega_0}(\omega_s)| e^{i\rho} \quad (1.B.4)$$

The real and imaginary parts of the Green's function are as follows:

$$\text{Re } G_{\omega_0}(\omega_s) = \frac{\omega_0^2 - \omega_s^2}{(\omega_0^2 - \omega_s^2)^2 + (2\Gamma\omega_s)^2}, \quad \text{Im } G_{\omega_0}(\omega_s) = \frac{2\Gamma\omega_s}{(\omega_0^2 - \omega_s^2)^2 + (2\Gamma\omega_s)^2}, \quad (1.B.5a) \quad (1.B.5b)$$

while its magnitude $|G(\omega_s)|$ and polar angle ρ are the following:

$$|G_{\omega_0}(\omega_s)| = \frac{1}{\sqrt{(\omega_0^2 - \omega_s^2)^2 + (2\Gamma\omega_s)^2}}, \quad \rho = \tan^{-1} \left(\frac{2\Gamma\omega_s}{\omega_0^2 - \omega_s^2} \right) \quad (1.B.5c) \quad (1.B.5d)$$

The angle ρ is the response *phase lag*, that is, the phase angle by which the response oscillation lags behind the phase $(-\omega_s t)$ of the stimulating oscillation.

$$z_{\text{response}}(t) = |G_{\omega_0}(\omega_s)| a(0) e^{-i(\omega_s t - \rho)} \quad (1.B.5)$$

It may help to visualize stimulus and response phasors as a pair rigidly rotating at rate ω_s . The response phasor lags ρ radians behind the stimulus as shown below in Fig. 1.B.2.

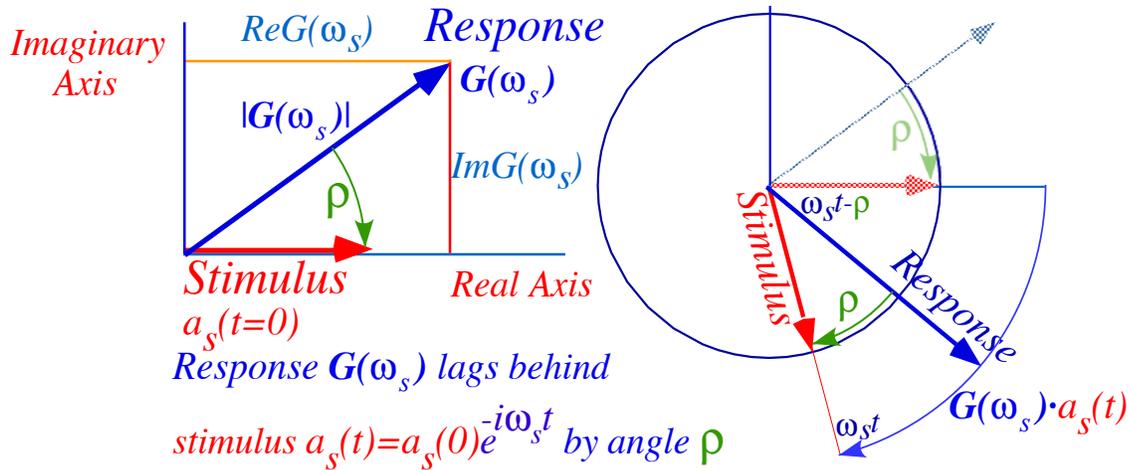


Fig.1.B.2 Oscillator response and stimulus phasors rotate rigidly at angular rate ω_s .

Complex algebra such as (1.B.5) is indispensable for analyzing oscillatory phenomena and devices. The most prevalent applications to electrical engineering (And, this is why *ABS*, $(xy) \rightarrow (polar)$, etc. keys are on your calculator!) began when Nikoli Tesla showed to Edison and Westinghouse (This took some time!) that their DC wiring was impractical and AC was the way to go in building a large power grid.

In optics as well as quantum mechanics there are many relations of the form (1.B.3) for which the *phasor* diagrams of Fig. 1.B.2 are helpful. Our description of quantum waves would be impossibly cumbersome without complex phasors and algebra.

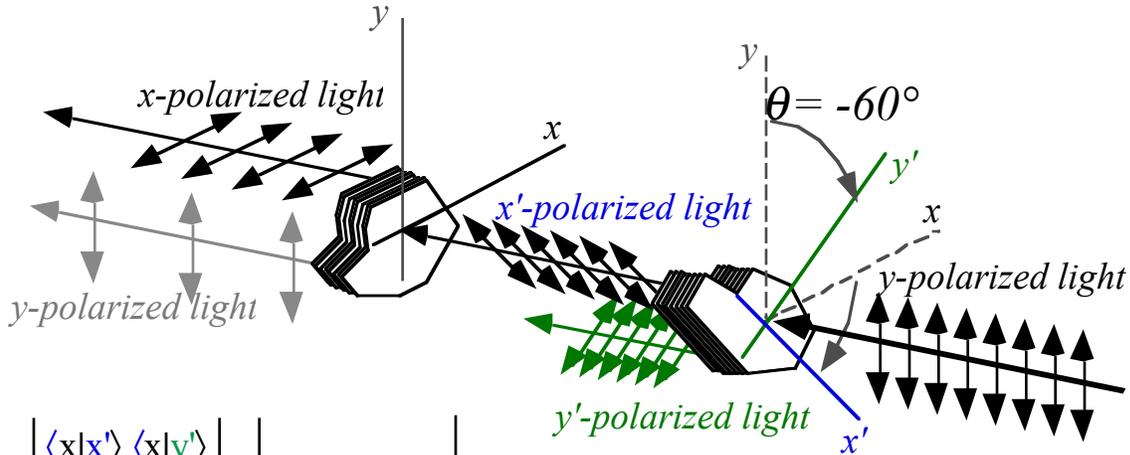
The idea of a phasor is based on that of *phase space*, certainly one of the oldest ideas of physics. Indeed, it goes back to the ancient astronomers tracking the phase of the moon. The real axis of a phasor is what we call the “is” of an oscillator, that is, where it is now. The imaginary component in Fig. 1.B.1 is what we will call the “gonna’be” component, that is, where the oscillator is going to be in $1/4$ -cycle if the phasor continues rotating clockwise at the same frequency. (The imaginary component is also a velocity or momentum component in ω units.) A helpful mnemonic is used in the billionaire business world:

Imagination precedes reality by one quarter!

In Unit 2 and throughout the rest of the text, phasors will be important tools.

Problems for Ch.1

1.2.1. A y-polarized light beam of unit amplitude (1 photon/sec.) enters the analyzer system as shown below. Fill in the blanks with numbers or symbols that tell as much as possible about what is present at each channel or branch.



$$\begin{vmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{vmatrix} = \begin{vmatrix} _ & _ \\ _ & _ \end{vmatrix}$$

$$\begin{vmatrix} \langle x'|x \rangle & \langle x'|y \rangle \\ \langle y'|x \rangle & \langle y'|y \rangle \end{vmatrix} = \begin{vmatrix} _ & _ \\ _ & _ \end{vmatrix}$$

State of output x channel

$|_ \rangle$
 Amplitude=
 Probability=

State of x' channel

$|_ \rangle$
 Amplitude=
 Probability=

State of input channel

$|y \rangle \frac{1.0}{1.0}$
 Amplitude=
 Probability=

State of output y channel

$|_ \rangle$
 Amplitude=
 Probability=

State of y' channel

$|_ \rangle$
 Amplitude=
 Probability=

A Dim View

1.2.2 (a) How far away from KUAF (10^5 Watts at 91.3 MHz) do you only get 1 photon/m²s?

(b) How far away from a 10^5 Watt green light source do you only get 1 photon/m²s? Can you guesstimate the threshold of visibility (in green photon/m²s) for your eye?

Give E-field amplitude in each case. Assume (incorrectly) scalar isotropic coherent wave sources.

Photonic Zeno

1.2.3 Imagine a series of N polarization beam sorters like the ones in Fig. 1.2.1 or 1.2.3 are placed so the top x-output beam of each goes into the next sorter in line which is rotated clockwise by an angle ϕ relative to the one before. Suppose unit amplitude x-polarization ($\Psi_x = 1, \Psi_y = 0$) comes into the first sorter in the series.

- What angle ϕ makes the amplitude $1/2^N$ coming out of this series? (Zeno attenuation)
- What angle ϕ makes the intensity $1/2^N$ coming out of this series? (Zeno depletion)
- Suppose the objective is to have as much y-polarization as is practical come out of this series. How does the output amplitude and intensity vary with the number N ?
How many (N) sorters are needed to give 99% photon conversion efficiency?

Electronic Zeno

1.2.4. Imagine a series of N electron beam sorters like the ones in Fig. 1.1.6 or 1.2.4 are placed so the top \uparrow - (up) output beam of each goes into the next sorter in line which is rotated clockwise by an angle ϕ relative to the one before. Suppose unit amplitude \uparrow - spin ($\Psi_{\uparrow} = 1, \Psi_{\downarrow} = 0$) comes into the first sorter in the series.

- What angle ϕ makes the amplitude $1/2^N$ coming out of this series? (Zeno attenuation)
- What angle ϕ makes the intensity $1/2^N$ coming out of this series? (Zeno depletion)
- Suppose the objective is to maximize \downarrow -spin (down) output from this series. How does the output amplitude and intensity vary with the number N ?
How many (N) sorters are needed to give 99% electron conversion efficiency?
(This is called adiabatic reversal.)

Fashion Plates

1.3.1. The effects of a 1/4-wave and a 1/2-wave plate were described in (1.3.1) to (1.3.3) and sketched in Fig. 1.3.6.

- Do the same for a "whole-wave" plate. (Give Ψ and sketch $\text{Re } \Psi$ trajectory.)
- Do the same for a 1/3-wave plate. (Give Ψ and sketch $\text{Re } \Psi$ trajectory.)

Good vibrations

1.B.1 An atomic oscillator is stimulated by a force that is sinusoidally oscillating at a frequency ω_S that is EQUAL to the natural resonance frequency ω_0 of the atom. Then the atom's oscillation will be...(a) in phase with (b) 45° ahead of (c) 45° behind (d) 90° ahead of (e) 90° behind (f) 135° ahead of (f) 135° behind (f) 180° out of phase with the phase of the force.

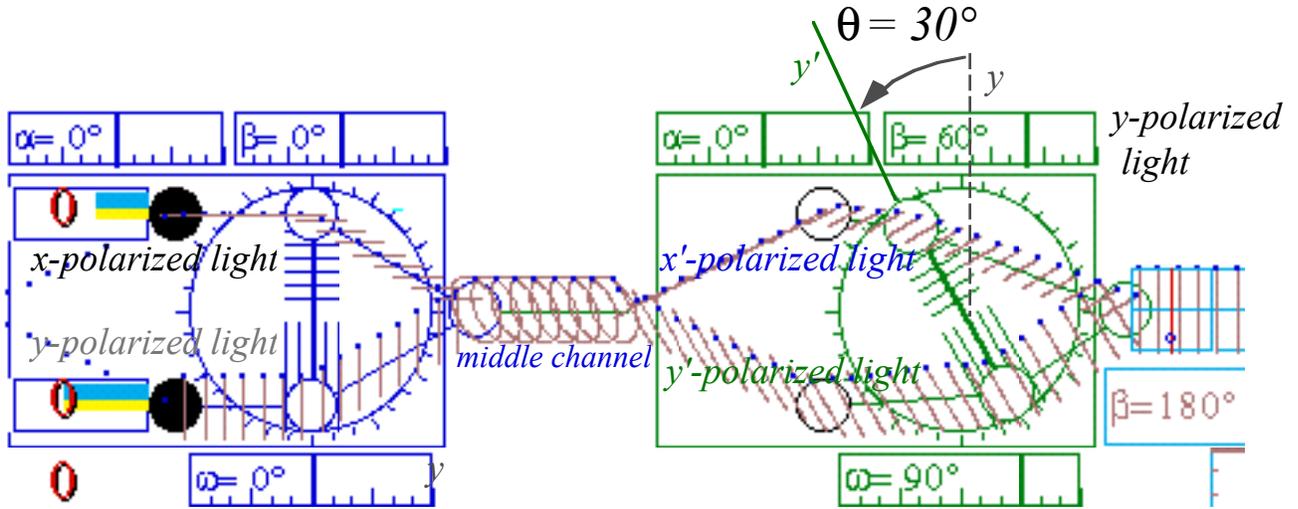
Keep the phase baby

1.B.1 An atomic oscillator is stimulated by a force sinusoidally oscillating at a frequency ω_S that is much LESS than the natural resonance frequency ω_0 of the atom. Then the atom's oscillation will be more or less...(a) in phase with (b) 45° ahead of (c) 45° behind (d) 90° ahead of (e) 90° behind (f) 135° ahead of (f) 135° behind (f) 180° out of phase with the phase of the force.

Ellipsis in Middle

1.3.1 A y-polarized light beam of unit amplitude (1 photon/sec.) enters an active analyzer that is tilted by 30° as shown below. The active analyzer puts a $\omega = 90^\circ$ phase factor $e^{-i\omega}$ in the x' beam.

Fill in the blanks with numbers or symbols that tell as much as possible about what is present at each channel or branch.



$$\begin{vmatrix} \langle x|x' \rangle & \langle x|y' \rangle \\ \langle y|x' \rangle & \langle y|y' \rangle \end{vmatrix} = \begin{vmatrix} _ & _ \\ _ & _ \end{vmatrix} \qquad \begin{vmatrix} \langle x'|x \rangle & \langle x'|y \rangle \\ \langle y'|x \rangle & \langle y'|y \rangle \end{vmatrix} = \begin{vmatrix} _ & _ \\ _ & _ \end{vmatrix}$$

State of output x channel

$|_ \rangle$
 Amplitude=
 Probability=

State of x' channel

$|_ \rangle$
 Amplitude=
 Probability=

State of input channel

$|y\rangle$ 1.0
 Amplitude=
1.0
 Probability=
1.0

State of middle channel

 Probability=

State of output y channel

$|_ \rangle$
 Amplitude=
 Probability=

State of y' channel

$|_ \rangle$
 Amplitude=
 Probability=

Note: The x' and y' quantities will not be graded. Use for work.

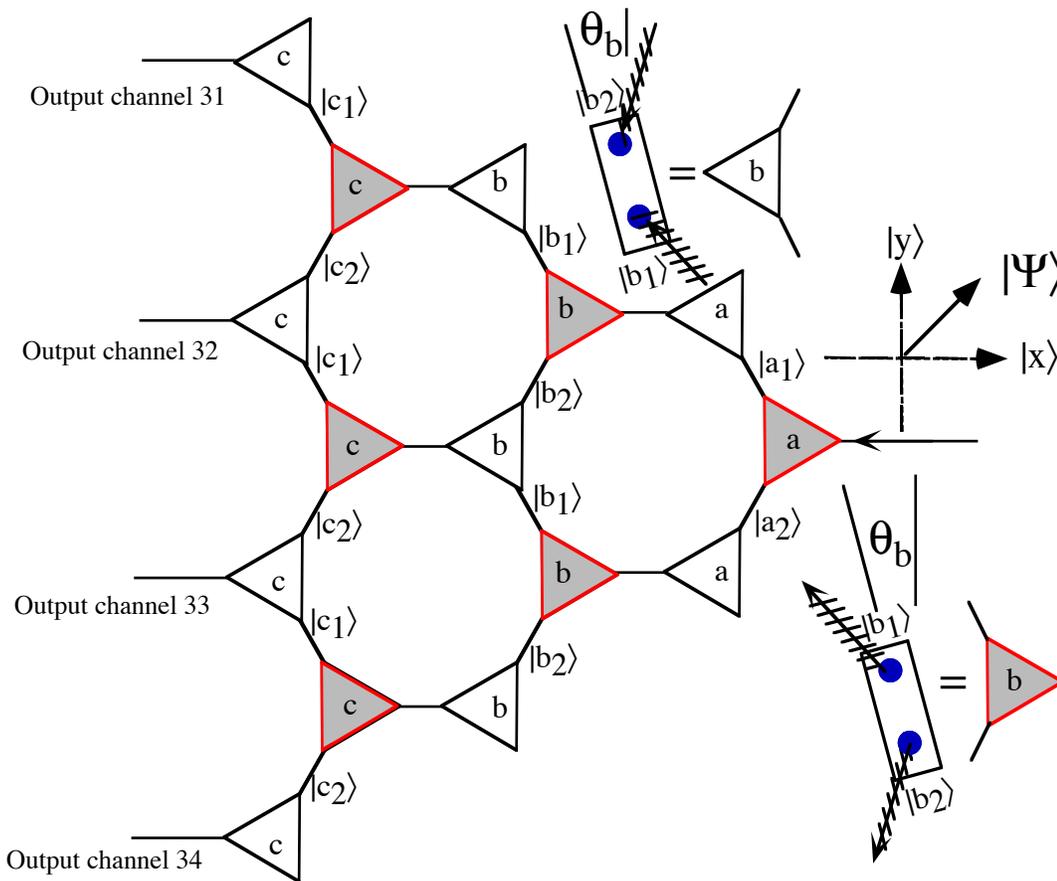
A Quantum Internet

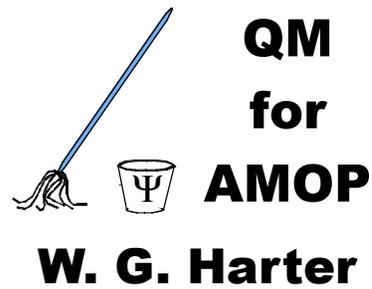
1.3.2. The sorter network shown in the third figure consists of three kinds of polarization sorters a, b, and c each with their main x-axis (a_1 -state) rotated by θ_a , θ_b , θ_c , respectively. The shaded triangles (pointing to the right toward the input side) split their input beams while the white triangles (pointing to the left) recombine the two beams of opposite polarization into one beam. (They are sorters reversed and flipped over.)

Let : $\theta_a = \theta$, $\theta_b = 2\theta$, $\theta_c = 3\theta$. Let input be x-polarized state $(\psi_x, \psi_y) = (1,0)$.

- (a) How many distinct "Feynman paths" are there from input to each output channel?
- (b) If the recombination is done with "peeking" or dephasing of some kind, compute the probability at each channel for the angle $\theta=45^\circ$.
- (c) If the recombination is done without "peeking" or dephasing of some kind, compute the probability at each channel for the angle $\theta=45^\circ$.
- (d) For each of the two cases of "peeking" and not "peeking" give a formula for probability in each channel as function of angle θ . (Check: Do your probabilities add up to one?)

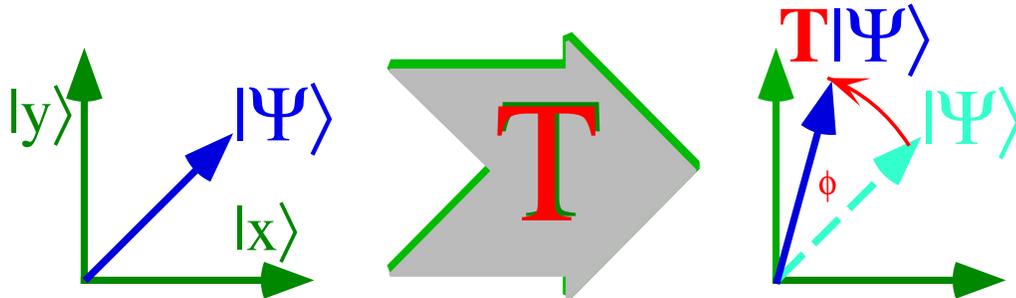
4. (Extra credit) A "peeking" analyzer gives the same probabilities in Fig. 1.3.8-9 whether it totally "collapses" the wave amplitudes or just jiggles one of their phases. Is there an experiment that could tell the difference? Tell why not, or describe one that would.





Chapter 2

Transformation and Transfer Operators and Matrices



Four axioms of quantum theory are based on the physics introduced in Chapter 1 and related to mathematical operations such as scalar products, matrix multiplication, change of basis, and projection. The four quantum axioms are related to the four axioms of a mathematical symmetry group as well as the all-important requirement of *unitarity* or “probability conservation.” The “mother of all groups,” the unitary group $U(N)$, is introduced for later use.

Chapter 2 Introduction to Transformation/Transfer Operators

- 2.1 Transformation Amplitude Matrices: Quantum Axioms45**
 - (a) Fundamental quantum axioms45
 - (1) The probability axiom45
 - (2) The conjugation or inversion axiom.....46
 - (3) The orthonormality or identity axiom46
 - (4) The completeness or closure axiom47
 - (b) Matrix bra-kets: bra-and-ket vectors and representations48
 - (1) Transformation matrix operation: Change of basis49
 - (2) Transformation matrix products: Unitary groups50
 - (3) Scalar products: Invariance and Hermitian conjugation51
 - (4) Particle expectation number: Norms and normalization52
 - (5) Axiom-4 totally abstracted: Projectors53
- 2.2 Transformation Operators: Unitarity and Group Axioms54**
 - (a) Base ket and bra transformations54
 - (b) Bra-ket vector component transformations55
 - (c) Group axioms56
 - (1) The closure axiom57
 - (2) The associativity axiom.....57
 - (3) The identity axiom.....57
 - (4) The inverse axiom58
 - (5) The commutative axiom (Abelian groups only)58
 - (d) U(n) group dimension.....58
 - (e) SU(n) group dimension.....58
- Problems for Ch. 259**

2.1 Transformation Amplitude Matrices: Quantum Axioms

We have seen how quantum amplitudes like $\langle x|x' \rangle$ are entries in arrays or *transformation matrices* like (1.2.5), for example. It is time to state mathematical and physical axioms which these quantities obey. This will help in reviewing the preceding sections and in establishing the mathematical basis of quantum theory.

The quantum axioms will be stated as clearly and as physically as possible. Like the establishment of a constitution for democratic country, it is imperative to base them on previous experience and above all make them reasonable, that is, *self-consistent*. In mathematics, there can be no proof of any part of a set of axioms. If such proof is found, (this happens rarely), then the axiom is upgraded to a theorem.

The same applies to physical theory, the main difference is that physical axioms in a given theory may eventually be "proved" by incorporation within a more fundamental theory. An example of this occurred when classical mechanics became superseded by quantum theory. Before quantum theory, the Newton's Laws such as conservation of momentum were physical axioms. Quantum theory (as we will see later on) "proves" that momentum is conserved, but only on the average and after many trials and counts. (Sound familiar? We have already seen that quantum theory, like a blow-dried weatherman, only predicts probabilities; the "hard-edged" classical world is gone forever.)

Let us now state a set of quantum axioms that seem to underlie the marvelous theory that has replaced the classical Newtonian theory which had reigned for nearly two centuries before 1913.

(a) Fundamental quantum axioms

Our statement of axioms will be based on an n -state system whose sorters always sort quantum beams into no more than n sub-beams. (Recall Fig. 1.1.1) Our motivation for the axioms will be based upon observed behavior for the 2-state systems of photon polarization. The axioms are concerned with the properties of n^2 complex quantum amplitudes $\langle j|k' \rangle$ arrayed in the following n -by- n *transformation matrix*.

$$T(b \leftarrow b') = \begin{pmatrix} \langle 1|1' \rangle & \langle 1|2' \rangle & \cdots & \langle 1|n' \rangle \\ \langle 2|1' \rangle & \langle 2|2' \rangle & \cdots & \langle 2|n' \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n|1' \rangle & \langle n|2' \rangle & \cdots & \langle n|n' \rangle \end{pmatrix} \quad (2.1.1)$$

(1) The probability axiom

Our first axiom concerns the physical interpretation of amplitudes.

Axiom 1: The absolute square $|\langle j|k' \rangle|^2 = \langle j|k' \rangle^ \langle j|k' \rangle$ gives the probability for occurrence in state- j of a system that started in state- $k'=1',2',\dots,$ or n' from one sorter and then was forced to choose between states $j=1,2,\dots,n$ by another sorter.*

This idea of probability was introduced first in Sect. 1.2b (See eq. (1.2.12).) The "forced-to-choose" clause and the word "state" are the kickers here. These concepts arise from the properties of "sorters" which have been described in the preceding sections. Some mysterious things are going on inside these sorters and analyzers so that a particle or system is forced to choose one and only one of n -states. That, in turn, makes the idea of a state mysterious, too. Which comes first, the chicken or the egg?

The next axiom is also a debatable one. Most quantum theories demand it, as will ours.

(2) The conjugation or inversion axiom

The second axiom concerns going backwards through a sorter or the reversal of amplitudes.

$$\text{Axiom 2: The complex conjugate of an amplitude gives its reverse: } \langle j|k \rangle^* = \langle k|j \rangle \quad (2.1.2)$$

We appeal to the idea of time-reversal and Planck's phase factor (1.2.7) when justifying this one.

Conjugation of $e^{-i\omega t}$ yields $(e^{-i\omega t})^* = e^{i\omega t}$ which can be interpreted as changing the sign of time ($t \rightarrow -t$), that is, going backwards. It also could have been a sign change of frequency ($\omega \rightarrow -\omega$), but the sign of ω is fixed by convention to be positive in elementary non-relativistic quantum theory. Interestingly, when it comes to doing a conjugation of the deBroglie phase factor e^{ikx} to $(e^{ikx})^* = e^{-ikx}$, we always interpret this as a change of sign of the wavevector ($k \rightarrow -k$) of momentum, that is, conjugation *really* makes things go backwards. Clearly, we haven't heard the last discussion of Axiom 2. Now an axiom that looks air-tight (but isn't).

(3) The orthonormality or identity axiom

The third axiom concerns the amplitude for "re measurement" by the same analyzer.

$$\text{Axiom 3: If identical analyzers are used twice or more the amplitude for a passed state-}k \text{ is one, and for all others it is forever zero: } \langle j|k \rangle = \delta_{jk} = \begin{cases} 1 \text{ if: } j=k \\ 0 \text{ if: } j \neq k \end{cases} = \langle j'|k' \rangle \quad (2.1.3)$$

You might think that a "caveman" could prove this axiom. We showed in Fig. 1.1.4 an easy experiment of repeated analysis of x -polarization. Apart from losses due to crystalline imperfections and absorption, this seems to "prove" unit probability for x -to- x . Probability, yes, but amplitude? Not necessarily! By axiom-1, probability is an absolute square which kills any phase factors an amplitude might have. The caveman could only prove Axiom-3 up to an arbitrary phase factor.

$$\langle j|k \rangle = e^{i\phi} \delta_{jk}$$

Unit phase is a mathematical convention we can live with. Now comes the "axiom of axioms."

(4) The completeness or closure axiom

The fourth axiom concerns the "Do-nothing" property of an ideal analyzer, that is, a sorter followed by an "unsorter" or "put-back-togetherer" as introduced in Sec. 1.3.

Axiom 4. Ideal sorting followed by ideal recombination of amplitudes has no effect:

$$\langle j|m' \rangle = \sum_{k=1}^n \langle j|k \rangle \langle k|m' \rangle \quad (2.1.4)$$

This axiom contains much if not all of the physical mystery of quantum phenomena. The idea is that a system initially in a prime state- m' is first sorted by an analyzer into states $k=1, 2, \dots, n$ each with amplitude

$\langle k | m' \rangle$ and then each of those is recombined (summed over k) and sorted by another analyzer into one of the doubly-prime states $j'' = 1'', 2'', \dots, n''$, each with an additional amplitude $\langle j'' | k \rangle$. The claim is that the amplitude for each of the doubly-prime j'' -states is none other than $\langle j'' | m' \rangle$, exactly what it would have been if k -analyzer had never intervened!

The upper halves of previous Figs. 1.3.8 and 1.3.9 show examples of this. How it can go wrong due to "peeking" or "dephasing" is shown by the lower halves of the same figures. To be precise, the figures show a case where the doubly-prime analyzer is the same kind of analyzer as the singly-prime sorter that would produce the initial state- m' , that is, $j'' = j'$. For this special case (2.1.4) becomes

$$\langle j' | m' \rangle = \sum_{k=1}^n \langle j' | k \rangle \langle k | m' \rangle \tag{2.1.5}$$

You can "verify" this case of Axiom-4. Using orthonormality Axiom-3 gives

$$\delta_{jm} = \sum_{k=1}^n \langle j' | k \rangle \langle k | m' \rangle = \sum_{k=1}^n \langle k | j' \rangle^* \langle k | m' \rangle \tag{2.1.6}$$

where conjugation axiom-2 is used, too. Now, for $m' = j'$ this becomes

$$1 = \sum_{k=1}^n \langle k | j' \rangle^* \langle k | j' \rangle = \sum_{k=1}^n |\langle k | j' \rangle|^2 = \sum_{k=1}^n P(j' \text{ to } k) \tag{2.1.7}$$

According to the probability axiom-1 this states that the sum of probabilities for all k -channels equals one which is consistent with the definition of probability in (1.1.2b). It means we have completely accounted for all possible states that a k -analyzer can sort. The k -states are then called a *complete set of states*.

A unit amplitude final x' -beam was shown in Fig. 1.3.9 along with a zero amplitude final y' -beam emerging from an ideal "do-nothing" analyzer with absolutely no "peeking" allowed. If the final $x'y'$ -sorter-counter had been replaced by an xy -analyzer or even a general Ψ -analyzer, it should still be impossible to see any effect of the intervening "do-nothing" xy -analyzer. This is the general physical consequence of Axiom-4. Again, precise "proof" of this or other axioms is currently impossible. They are just some "laws" that we must live with for awhile...probably, quite awhile.

(b) Matrix bra-kets: bra-and-ket vectors and representations

The quantum axioms are strongly connected to mathematical axioms of linear algebra and unitary matrix group theory. As mentioned before in Secs. 1.2a and c, *Dirac notation* for vectors is a result of "dissecting" a transformation matrix. Dirac invented a notation for entire columns and entire rows of a T -matrix; they are called *kets* $|k\rangle$ and *bras* $\langle j|$ respectively. The general T -matrix below shows its bras and kets.

$$T(b \leftarrow b') = \begin{matrix} \begin{matrix} |1'\rangle & |2'\rangle & \dots & |n'\rangle \\ \downarrow & \downarrow & \dots & \downarrow \end{matrix} \\ \left(\begin{matrix} \langle 1|1'\rangle & \langle 1|2'\rangle & \dots & \langle 1|n'\rangle \\ \langle 2|1'\rangle & \langle 2|2'\rangle & \dots & \langle 2|n'\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n|1'\rangle & \langle n|2'\rangle & \dots & \langle n|n'\rangle \end{matrix} \right) \leftarrow \begin{matrix} \langle 1| \\ \langle 2| \\ \vdots \\ \langle n| \end{matrix} \end{matrix} \tag{2.1.8}$$

These bras and kets are examples of what we will call *abstract* mathematical quantities. They stand for a definite physical object, in this case, definite states occupied by particles or systems. In order to view these objects or store them in a computer, you need a *representation* of them, that is, a list of numbers such as the *column vectors* representing kets which stand for primed states $1', 2'$, as pointed out in (2.1.8) or below,

$$|1'\rangle \rightarrow \begin{pmatrix} \langle 1|1'\rangle \\ \langle 2|1'\rangle \\ \vdots \\ \langle n|1'\rangle \end{pmatrix}, \quad |2'\rangle \rightarrow \begin{pmatrix} \langle 1|2'\rangle \\ \langle 2|2'\rangle \\ \vdots \\ \langle n|2'\rangle \end{pmatrix}, \quad (2.1.9a)$$

or the *row vectors* representing bras standing for unprimed states-1, 2,..as pointed out below.

$$\langle 1| \rightarrow \left(\langle 1|1'\rangle \quad \langle 1|2'\rangle \quad \dots \quad \langle 1|n'\rangle \right), \quad \langle 2| \rightarrow \left(\langle 2|1'\rangle \quad \langle 2|2'\rangle \quad \dots \quad \langle 2|n'\rangle \right) \quad (2.1.9b)$$

This Dirac abstraction is achieved by literally "abstracting" bras or kets from Axiom-4. Starting with an Axiom-4 equation for any basis $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$ or $\{|\bar{1}\rangle, |\bar{2}\rangle, \dots, |\bar{n}\rangle\}$ or $\{|1'\rangle, |2'\rangle, \dots, |n'\rangle\}$, etc.

$$\langle m''|j'\rangle = \sum_{k=1}^n \langle m''|k\rangle \langle k|j'\rangle = \sum_{\bar{k}=1}^n \langle m''|\bar{k}\rangle \langle \bar{k}|j'\rangle = \sum_{k'=1}^n \langle m''|k'\rangle \langle k'|j'\rangle = \text{etc.}, \quad (2.1.10a)$$

one can "rip-off" its bra $\langle m''|$ to give a ket $|j'\rangle$

$$|j'\rangle = \sum_{k=1}^n |k\rangle \langle k|j'\rangle = \sum_{\bar{k}=1}^n |\bar{k}\rangle \langle \bar{k}|j'\rangle = \sum_{k'=1}^n |k'\rangle \langle k'|j'\rangle = \text{etc.}, \quad (2.1.10b)$$

or the j' -ket can be "ripped-off" to expose a bra. (Recall the finale of 1999 US World Soccer Cup victory.)

$$\langle m''| = \sum_{k=1}^n \langle m''|k\rangle \langle k| = \sum_{\bar{k}=1}^n \langle m''|\bar{k}\rangle \langle \bar{k}| = \sum_{k'=1}^n \langle m''|k'\rangle \langle k'| = \text{etc.} \quad (2.1.10c)$$

A representation depends upon the basis you are using. Each case of the (2.1.10b) uses different transformation matrix coefficients and base kets to describe a single abstract ket. The following are all different representations of the same ket $|1'\rangle$. The third one is the j' -representation of $|1'\rangle$ which is composed of all zero components but one according to Axiom-3. (It's being represented in its own basis.)

$$|1'\rangle \rightarrow \begin{pmatrix} \langle 1|1'\rangle \\ \langle 2|1'\rangle \\ \vdots \\ \langle n|1'\rangle \end{pmatrix}, \text{ or: } |1'\rangle \rightarrow \begin{pmatrix} \langle \bar{1}|1'\rangle \\ \langle \bar{2}|1'\rangle \\ \vdots \\ \langle \bar{n}|1'\rangle \end{pmatrix}, \text{ or: } |1'\rangle \rightarrow \begin{pmatrix} \langle 1'|1'\rangle \\ \langle 2'|1'\rangle \\ \vdots \\ \langle n'|1'\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ etc.}$$

The same applies to each of the bra representations (2.1.10c). Here are some representations of bra $\langle 2'|$.

$$\langle 2'| \rightarrow \left(\langle 2'|1\rangle \quad \langle 2'|2\rangle \quad \dots \quad \langle 2'|n\rangle \right), \text{ or: } \langle 2'| \rightarrow \left(\langle 2'|\bar{1}\rangle \quad \langle 2'|\bar{2}\rangle \quad \dots \quad \langle 2'|\bar{n}\rangle \right), \\ \text{ or: } \langle 2'| \rightarrow \left(\langle 2'|1'\rangle \quad \langle 2'|2'\rangle \quad \dots \quad \langle 2'|n'\rangle \right) = \left(0 \quad 1 \quad \dots \quad 0 \right), \text{ etc.}$$

The last is the j' -representation of $\langle 2'|$ using its own (primed) basis.

Note: It is very tempting to replace the "mapping arrows" (\rightarrow) by equal signs ($=$). This is often done in many a sloppy modern physics text (including this one, later on!), but don't do it until you are sure of which

basis you are using and are settled into using that basis exclusively. Once one is "married" to a given basis it is permissible to dispense with some formalities!

Abstract forms (2.1.10b-c) can be reassembled with a general state-ket $|\Psi\rangle$ or bra $\langle\Psi|$ to give the following two Dirac forms of Axiom-4 for an arbitrary state- Ψ .

$$\langle m''|\Psi\rangle = \sum_{k=1}^n \langle m''|k\rangle\langle k|\Psi\rangle = \sum_{\bar{k}=1}^n \langle m''|\bar{k}\rangle\langle\bar{k}|\Psi\rangle = \sum_{k'=1}^n \langle m''|k'\rangle\langle k'|\Psi\rangle = etc. , \tag{2.1.11a}$$

$$\langle\Psi|j'\rangle = \sum_{k=1}^n \langle\Psi|k\rangle\langle k|j'\rangle = \sum_{\bar{k}=1}^n \langle\Psi|\bar{k}\rangle\langle\bar{k}|j'\rangle = \sum_{k'=1}^n \langle\Psi|k'\rangle\langle k'|j'\rangle = etc. \tag{2.1.11b}$$

(1) Transformation matrix operation: Change of basis

Axiom-4 has an important mathematical interpretation. You may view it as a *matrix-vector transformation* which effects a *change of basis*. Compare the Dirac algebraic form (2.1.11a) to the following matrix-acting-on-column-vector multiplication.

$$\begin{pmatrix} \langle 1''|\Psi\rangle \\ \langle 2''|\Psi\rangle \\ \vdots \\ \langle n''|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \langle 1''|1'\rangle & \langle 1''|2'\rangle & \dots & \langle 1''|n'\rangle \\ \langle 2''|1'\rangle & \langle 2''|2'\rangle & \dots & \langle 2''|n'\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n''|1'\rangle & \langle n''|2'\rangle & \dots & \langle n''|n'\rangle \end{pmatrix} \begin{pmatrix} \langle 1'|\Psi\rangle \\ \langle 2'|\Psi\rangle \\ \vdots \\ \langle n'|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \langle 1''|\bar{1}\rangle & \langle 1''|\bar{2}\rangle & \dots & \langle 1''|\bar{n}\rangle \\ \langle 2''|\bar{1}\rangle & \langle 2''|\bar{2}\rangle & \dots & \langle 2''|\bar{n}\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n''|\bar{1}\rangle & \langle n''|\bar{2}\rangle & \dots & \langle n''|\bar{n}\rangle \end{pmatrix} \begin{pmatrix} \langle \bar{1}|\Psi\rangle \\ \langle \bar{2}|\Psi\rangle \\ \vdots \\ \langle \bar{n}|\Psi\rangle \end{pmatrix} = etc. \tag{2.1.11a}_{representation}$$

A 2-by-2 example of this was first given by (1.2.6b). The Dirac algebraic form for the bra expression (2.1.11b) has the following row-vector-on-matrix representation.

$$\begin{pmatrix} \langle\Psi|1'\rangle & \langle\Psi|2'\rangle & \dots & \langle\Psi|n'\rangle \end{pmatrix} = \begin{pmatrix} \langle\Psi|1''\rangle & \langle\Psi|2''\rangle & \dots & \langle\Psi|n''\rangle \end{pmatrix} \begin{pmatrix} \langle 1''|1'\rangle & \langle 1''|2'\rangle & \dots & \langle 1''|n'\rangle \\ \langle 2''|1'\rangle & \langle 2''|2'\rangle & \dots & \langle 2''|n'\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n''|1'\rangle & \langle n''|2'\rangle & \dots & \langle n''|n'\rangle \end{pmatrix} = etc. \tag{2.1.11b}_{representation}$$

These are the bra and ket forms of general quantum coordinate transformations.

(2) Transformation matrix products: Unitary groups

Axiom-4 has another even more important mathematical interpretation. You may view it as a *transformation matrix product* which forms a *unitary transformation group U(n)*. Axiom-4 is basically a matrix product as seen by comparing the following two representations. First, the original Dirac form

$$\langle j''|m'\rangle = \sum_{k=1}^n \langle j''|k\rangle\langle k|m'\rangle \tag{2.1.12a}$$

and then here is the same thing in matrix form.

$$\begin{pmatrix} \langle 1''|1'\rangle & \langle 1''|2'\rangle & \dots & \langle 1''|n'\rangle \\ \langle 2''|1'\rangle & \langle 2''|2'\rangle & \dots & \langle 2''|n'\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n''|1'\rangle & \langle n''|2'\rangle & \dots & \langle n''|n'\rangle \end{pmatrix} = \begin{pmatrix} \langle 1''|1\rangle & \langle 1''|2\rangle & \dots & \langle 1''|n\rangle \\ \langle 2''|1\rangle & \langle 2''|2\rangle & \dots & \langle 2''|n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n''|1\rangle & \langle n''|2\rangle & \dots & \langle n''|n\rangle \end{pmatrix} \bullet \begin{pmatrix} \langle 1|1'\rangle & \langle 1|2'\rangle & \dots & \langle 1|n'\rangle \\ \langle 2|1'\rangle & \langle 2|2'\rangle & \dots & \langle 2|n'\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n|1'\rangle & \langle n|2'\rangle & \dots & \langle n|n'\rangle \end{pmatrix} \tag{2.1.12b}$$

You could also write this is as the following *matrix product*

$$T_{j''m'} \begin{pmatrix} \text{prime} \\ \text{to} \\ \text{double-prime} \end{pmatrix} = \sum_{k=1}^n T_{j''k} \begin{pmatrix} \text{unprimed} \\ \text{to} \\ \text{double-prime} \end{pmatrix} T_{k m'} \begin{pmatrix} \text{prime} \\ \text{to} \\ \text{unprimed} \end{pmatrix} \quad (2.1.12c)$$

or abstractly as follows. $\mathbf{T}(b'' \leftarrow b') = \mathbf{T}(b'' \leftarrow b) \bullet \mathbf{T}(b \leftarrow b')$ (2.1.12d)

The latter **T**'s are our first example of *abstract group operators*. They are quantum coordinate transformation operators that deliver us from one basis to another through transformations such as (2.1.11). Eq. (2.1.12) is their composition rule: If **T**(B ← A) goes from A to B and **T**(C ← B) goes from B to C then the product **T**(C ← B)•**T**(B ← A) goes directly from A to C, that is, it equals **T**(C ← A) .

These operators stand for all the possible transformations that are allowed in quantum space. Their matrix representations represent all the means for "getting around" in the state space. As we will see later, they also represent (practically) all the possible quantum analyzers that can be built in laboratory experiments involving the n-state system being studied. Remember that each abstract mathematical quantity corresponds to some "thing" out there. Bras and kets correspond to systems of particles and operators like the **T**'s correspond to devices which "do things" to the systems, that is, transport or "taxi" the system from one state to another.

(3) Scalar products: Invariance and Hermitian conjugation

Finally, Axiom-4 has a very simple mathematical interpretation. You may view it as a *scalar product* between two arbitrary state vectors, a state-ket $|\Psi\rangle$ and bra $\langle\Phi|$. First, the general Axiom-4

$$\langle\Phi|\Psi\rangle = \sum_{k=1}^n \langle\Phi|k\rangle\langle k|\Psi\rangle = \sum_{\bar{k}=1}^n \langle\Phi|\bar{k}\rangle\langle\bar{k}|\Psi\rangle = etc. \quad (2.1.13a)$$

and then the representations of Axiom-4 as a "dot" product of row and column vectors.

$$\langle\Phi|\Psi\rangle = \left(\langle\Phi|1\rangle \ \langle\Phi|2\rangle \ \dots \ \langle\Phi|n\rangle \right) \begin{pmatrix} \langle 1|\Psi\rangle \\ \langle 2|\Psi\rangle \\ \vdots \\ \langle n|\Psi\rangle \end{pmatrix} = \left(\langle\Phi|\bar{1}\rangle \ \langle\Phi|\bar{2}\rangle \ \dots \ \langle\Phi|\bar{n}\rangle \right) \begin{pmatrix} \langle \bar{1}|\Psi\rangle \\ \langle \bar{2}|\Psi\rangle \\ \vdots \\ \langle \bar{n}|\Psi\rangle \end{pmatrix} = etc. \quad (2.1.13b)$$

In Sec. b(1-2) we saw matrix multiplication operations (2.1.11) and (2.1.12). Note that these are just combinations of the elementary scalar product shown in (2.1.13), that is, a row "dot-multiplied" with a column. Matrix operation on a vector in (2.1.11) involves *n* scalar products. Matrix multiplication (2.1.12) involves *n*² scalar products. It helps to see how Axiom-4 relates all these operations.

The scalar product $\langle\Phi|\Psi\rangle$ is one of few abstract quantum quantities that has the same numerical value in all possible representations. All the possible sums in (2.1.13) give the same number $\langle\Phi|\Psi\rangle$ even though the various representations of individual abstract bra $\langle\Phi|$ and ket $|\Psi\rangle$ will have wildly different numbers in them. The scalar products are said to be *invariant to change of bases*. This invariance is called *unitary invariance* for reasons that will be discussed shortly.

The scalar product can be expressed entirely in terms of ket components by using the conjugation axiom-2 to convert bra components $\langle\Phi|k\rangle$ to conjugated ket components $\langle k|\Phi\rangle^* = \langle\Phi|k\rangle$.

$$\langle \Phi | \Psi \rangle = \sum_{k=1}^n \langle k | \Phi \rangle^* \langle k | \Psi \rangle = \sum_{\bar{k}=1}^n \langle \bar{k} | \Phi \rangle^* \langle \bar{k} | \Psi \rangle = \text{etc.} \quad (2.1.14)$$

$$\langle \Phi | \Psi \rangle = \begin{pmatrix} \langle 1 | \Phi \rangle^* & \langle 2 | \Phi \rangle^* & \dots & \langle n | \Phi \rangle^* \end{pmatrix} \begin{pmatrix} \langle 1 | \Psi \rangle \\ \langle 2 | \Psi \rangle \\ \vdots \\ \langle n | \Psi \rangle \end{pmatrix} = \begin{pmatrix} \langle \bar{1} | \Phi \rangle^* & \langle \bar{2} | \Phi \rangle^* & \dots & \langle \bar{n} | \Phi \rangle^* \end{pmatrix} \begin{pmatrix} \langle \bar{1} | \Psi \rangle \\ \langle \bar{2} | \Psi \rangle \\ \vdots \\ \langle \bar{n} | \Psi \rangle \end{pmatrix} = \text{etc.}$$

The mathematical "sex change" operation of converting a ket to a bra is given the technical name of *transpose conjugation* †. (Perhaps, the dagger symbol † indicates the use of a knife or scapel.) It is also called *Hermitian conjugation* after the mathematician Hermite. (No, Hermite did not undergo any such operation, so far as we know.) As you can see from (2.1.14), this operation on column vectors is relatively painless and, indeed, receives far less publicity than its human equivalent. It simply involves the flipping of a column to become a row or vice-versa (called *transposing* ^T) followed or preceded by complex conjugation (*) of every amplitude in the row or column.

The notation for Hermitian conjugation is used as follows. First the abstract form is as follows.

$$\langle \Phi |^\dagger = | \Phi \rangle, \text{ or: } | \Phi \rangle^\dagger = \langle \Phi | \quad (2.1.15)$$

A matrix representation of this *conjugation* † is a row-to-column conversion with complex conjugation.

$$\begin{pmatrix} \langle \Phi | 1 \rangle & \langle \Phi | 2 \rangle & \dots & \langle \Phi | n \rangle \end{pmatrix}^\dagger = \begin{pmatrix} \langle \Phi | 1 \rangle^* \\ \langle \Phi | 2 \rangle^* \\ \vdots \\ \langle \Phi | n \rangle^* \end{pmatrix} = \begin{pmatrix} \langle 1 | \Phi \rangle \\ \langle 2 | \Phi \rangle \\ \vdots \\ \langle n | \Phi \rangle \end{pmatrix}, \text{ or: } \begin{pmatrix} \langle 1 | \Phi \rangle \\ \langle 2 | \Phi \rangle \\ \vdots \\ \langle n | \Phi \rangle \end{pmatrix}^\dagger = \begin{pmatrix} \langle 1 | \Phi \rangle^* & \langle 2 | \Phi \rangle^* & \dots & \langle n | \Phi \rangle^* \end{pmatrix} \\ = \begin{pmatrix} \langle \Phi | 1 \rangle & \langle \Phi | 2 \rangle & \dots & \langle \Phi | n \rangle \end{pmatrix}$$

Axiom-2 was used, again. Note: two Hermitian conjugations cancel: $\mathbf{A}^{\dagger\dagger} = \mathbf{A}$. (Don't try this at home!)

(4) Particle expectation number: Norms and normalization

A most important physical quantity is the scalar product of a state vector $|\Psi\rangle$ with itself, conjugated, of course. (You've probably noticed that "same-sex" marriages are prohibited in this quantum state theory. Actually, it can be done, but they're not real. Or invariant. In fact, they're usually very complex. For these you must wait until we discuss such taboo topics as *n-body mechanics* and *tensor operator theory*. Can't wait, can you?) The self product of $|\Psi\rangle$ with itself (which is always real) is quite analogous to scalar product.

$$\langle \Psi | \Psi \rangle = \sum_{k=1}^n \langle \Psi | k \rangle \langle k | \Psi \rangle = \sum_{k=1}^n \langle k | \Psi \rangle^* \langle k | \Psi \rangle = \sum_{k=1}^n |\langle k | \Psi \rangle|^2 = \begin{pmatrix} \langle 1 | \Psi \rangle^* & \langle 2 | \Psi \rangle^* & \dots & \langle n | \Psi \rangle^* \end{pmatrix} \begin{pmatrix} \langle 1 | \Psi \rangle \\ \langle 2 | \Psi \rangle \\ \vdots \\ \langle n | \Psi \rangle \end{pmatrix} \quad (2.1.16)$$

$\langle \Psi | \Psi \rangle$ is called the *particle number expectation* or *total probability* for a general quantum state $|\Psi\rangle$.

According to Axiom-1, the probability axiom, $\langle \Psi | \Psi \rangle$ is just the sum of the channel probabilities $|\langle k | \Psi \rangle|^2$ for each k-channel. Initially, we expect the probability sum to be unity as in (1.12b) or (2.1.7). However, after a particle or system is dragged through several analyzers having counters and other such dissipative devices, the total probability may be reduced significantly. An example was the x-beam after "peeking" which had a state vector $|\Psi\rangle = \begin{pmatrix} \sqrt{0.75} & 0 \end{pmatrix}$ for which the total probability is only 75%.

$\langle \Psi | \Psi \rangle$ is also called a *state norm*. For "permissive" states, the norm can be any real number. However, states that are being used as *base states* are expected to set a better example. Base states are supposed to be a complete set of kets $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$ and bras $\{\langle 1|, \langle 2|, \dots, \langle n|\}$ which satisfy all Axioms 1-4. In particular, they must be normal and satisfy the *ortho-normalization conditions* of Axiom-3.

$$\langle i | j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if: } i=j \\ 0 & \text{if: } i \neq j \end{cases} \tag{2.1.17}$$

The *Kronecker delta function* δ_{ij} is intended to guarantee unit norm $\langle j | j \rangle = 1$ (for all j) and, hence, unit probability for all base states. But, state vectors become "sub-normal" or abnormal in "dissipative" places.(5)

Axiom-4 totally abstracted: Projectors

Axiom-4 has one more level of abstraction: a complete "rip-off" of its bras and kets as follows.

$$\langle j | m \rangle = \sum_{k=1}^n \langle j | k \rangle \langle k | m \rangle \Rightarrow \mathbf{1} = \sum_{k=1}^n |k\rangle \langle k| \tag{2.1.18}$$

The result is a sum of "ket-bras" $|k\rangle \langle k|$ which are quite the opposite of the "bra-kets" like $\langle k | k \rangle$ which are scalars. (By axiom-3 $\langle k | k \rangle = 1$.) The ket-bras $|k\rangle \langle k|$ are elementary examples of *tensor operators* and are called *projection operators* or *projectors* \mathbf{P}_k .

$$\mathbf{P}_k = |k\rangle \langle k| \tag{2.1.19}$$

The axiom-4 sum to the identity operator $\mathbf{1}$ is called a *completeness relation*.

$$\mathbf{1} = \sum_{k=1}^n \mathbf{P}_k = \sum_{k=1}^n |k\rangle \langle k| \tag{2.1.20}$$

The action of a projection operator \mathbf{P}_k on a general state vector $|\Psi\rangle$ yields a projection or "shadow" $|k\rangle \langle k | \Psi \rangle$ of the original vector in the direction of the k-th base ket, as shown in the Fig. 2.1.1.

$$\mathbf{P}_k |\Psi\rangle = |k\rangle \langle k | \Psi \rangle \tag{2.1.21}$$

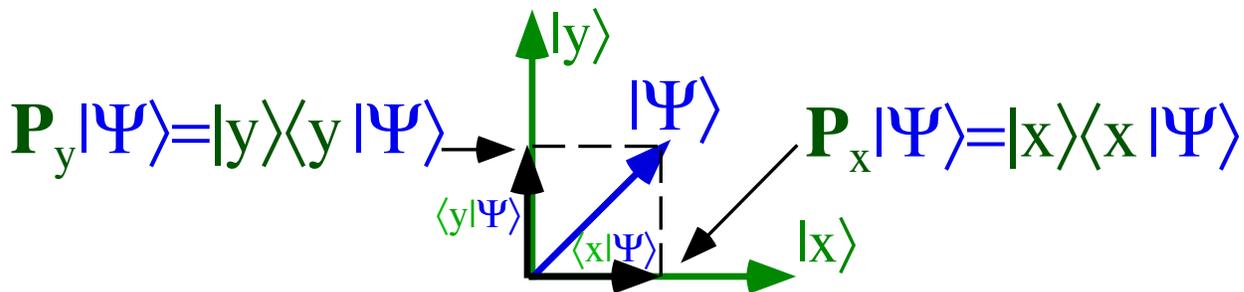


Fig. 2.1.1 Projection operators \mathbf{P}_k map state ket onto base axes

To construct a projection matrix representation of a \mathbf{P}_k simply reattach a \mathbf{P}_k 's "ripped-off" bras and kets for the desired basis, say the ϕ -tilted polarization bases $\{|x'\rangle, |y'\rangle\}$.

$$\begin{pmatrix} \langle x' | \mathbf{P}_x | x' \rangle & \langle x' | \mathbf{P}_x | y' \rangle \\ \langle y' | \mathbf{P}_x | x' \rangle & \langle y' | \mathbf{P}_x | y' \rangle \end{pmatrix} = \begin{pmatrix} \langle x' | x \rangle \langle x | x' \rangle & \langle x' | x \rangle \langle x | y' \rangle \\ \langle y' | x \rangle \langle x | x' \rangle & \langle y' | x \rangle \langle x | y' \rangle \end{pmatrix}, \text{ where: } \begin{pmatrix} \langle x | x' \rangle & \langle x | y' \rangle \\ \langle y | x' \rangle & \langle y | y' \rangle \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

Projector $\mathbf{P}_x = |x\rangle\langle x|$ is what is called an *outer* or *Kronecker tensor* (\otimes) *product* of ket and bra .

$$\begin{pmatrix} \langle x' | \mathbf{P}_x | x' \rangle & \langle x' | \mathbf{P}_x | y' \rangle \\ \langle y' | \mathbf{P}_x | x' \rangle & \langle y' | \mathbf{P}_x | y' \rangle \end{pmatrix} = \begin{pmatrix} \langle x' | x \rangle \langle x | x' \rangle & \langle x' | x \rangle \langle x | y' \rangle \\ \langle y' | x \rangle \langle x | x' \rangle & \langle y' | x \rangle \langle x | y' \rangle \end{pmatrix} = \begin{pmatrix} \langle x' | x \rangle \\ \langle y' | x \rangle \end{pmatrix} \otimes \begin{pmatrix} \langle x | x' \rangle & \langle x | y' \rangle \end{pmatrix} \quad (2.1.22)$$

The $x'y'$ -representations for both \mathbf{P}_x and \mathbf{P}_y are worked out below.

$$\begin{aligned} \mathbf{P}_x = |x\rangle\langle x| &\rightarrow \begin{pmatrix} \cos \phi \\ -\sin \phi \end{pmatrix} \otimes \begin{pmatrix} \cos \phi & -\sin \phi \end{pmatrix} & \mathbf{P}_y = |y\rangle\langle y| &\rightarrow \begin{pmatrix} \sin \phi \\ \cos \phi \end{pmatrix} \otimes \begin{pmatrix} \sin \phi & \cos \phi \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \phi & -\sin \phi \cos \phi \\ -\sin \phi \cos \phi & \sin^2 \phi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_{(\text{for } \phi=0)}, & & = \begin{pmatrix} \sin^2 \phi & \sin \phi \cos \phi \\ \sin \phi \cos \phi & \cos^2 \phi \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_{(\text{for } \phi=0)} \end{aligned}$$

The case $\phi=0$ gives xy -representations of \mathbf{P}_x and \mathbf{P}_y which each contain all "0" except for a single "1". Such matrices are called *elementary unit tensor representations* or *unit dyads*.

2.2 Transformation Operators: Unitarity and Group Axioms

If there is one thing you should always remember about quantum operators or any sort of mathematical operator, it is this. *Always begin with a base-state definition of an operator.* This simple rule will save you endless headache and hair-pulling.

(a) Base ket and bra transformations

As an example, consider a rotational transformation $\mathbf{T}=\mathbf{R}(\phi)$ similar to the one used for the rotated polarizer in Sec. 1.2. You want a transformation operator that takes each Cartesian ket $\{|x\rangle, |y\rangle\}$ and maps it into a new rotated basis $\{|\bar{x}\rangle = \mathbf{T}|x\rangle, |\bar{y}\rangle = \mathbf{T}|y\rangle\}$ given by the following and shown in Fig. 2.2.1.

$$\{|\bar{x}\rangle = \mathbf{T}|x\rangle = \cos\phi|x\rangle + \sin\phi|y\rangle, \quad |\bar{y}\rangle = \mathbf{T}|y\rangle = -\sin\phi|x\rangle + \cos\phi|y\rangle\} \quad (2.2.1)$$

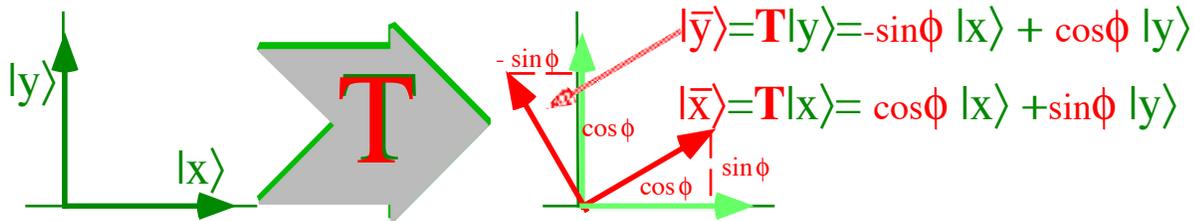


Fig. 2.2.1 Transformation \mathbf{T} maps unit kets into rotated ones

Once this basic ket-vector definition \mathbf{T} is given, all else follows. Scalar products of (2.2.1) with bra $\langle x|$ and then by bra $\langle y|$ gives the four transformation matrix components. Because of axiom-2 orthonormality $\langle j|k\rangle = \delta_{jk}$ or $\langle x|x\rangle = 1$, $\langle x|y\rangle = 0$, etc., each product $\langle j|\bar{k}\rangle$ gives just one term.

$$\begin{pmatrix} \langle x|\bar{x}\rangle & \langle x|\bar{y}\rangle \\ \langle y|\bar{x}\rangle & \langle y|\bar{y}\rangle \end{pmatrix} = \begin{pmatrix} \langle x|\mathbf{T}|x\rangle & \langle x|\mathbf{T}|y\rangle \\ \langle y|\mathbf{T}|x\rangle & \langle y|\mathbf{T}|y\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \quad (2.2.2)$$

The general n-state form of transformation (2.2.1) is

$$|\bar{k}\rangle = \mathbf{T}|k\rangle = \sum_{j=1}^n |j\rangle\langle j|\bar{k}\rangle = \sum_{j=1}^n |j\rangle\langle j|\mathbf{T}|k\rangle, \quad \text{where: } \langle j|\bar{k}\rangle = \langle j|\mathbf{T}|k\rangle \quad (2.2.3)$$

Axiom-2 requires that all base states satisfy orthonormality, and so must the new $\{|\bar{x}\rangle, |\bar{y}\rangle\}$ bases.

$$\langle j|k\rangle = \delta_{jk} = \langle \bar{j}|\bar{k}\rangle$$

In order for the transformation \mathbf{T} to preserve orthonormality, the operator which transforms old bras $\{\langle x|, \langle y|\}$ into the new bras $\{\langle \bar{x}|, \langle \bar{y}|\}$ must be an inverse \mathbf{T}^{-1} to the transformation \mathbf{T} that transformed the kets in the basic definition (2.2.1). Only then does a scalar product stay the same. That is,

$$\langle \bar{x}| = \langle x|\mathbf{T}^{-1}, \quad \langle \bar{y}| = \langle y|\mathbf{T}^{-1} \quad \text{so that: } \langle \bar{j}|\bar{k}\rangle = \langle j|\mathbf{T}^{-1}\mathbf{T}|k\rangle = \langle j|\mathbf{1}|k\rangle = \langle j|k\rangle. \quad (2.2.4)$$

However, bras are obtained by Hermitian conjugation according to (2.1.15)

$$\langle \bar{k}| = |\bar{k}\rangle^\dagger = (\mathbf{T}|k\rangle)^\dagger = \left(\sum_{j=1}^n |j\rangle\langle j|\bar{k}\rangle \right)^\dagger$$

How do we dagger (\dagger) a whole bunch of terms? The answer, is to \dagger -stab (Hermite-transpose) each part.

$$\langle \bar{k} | = |k\rangle^\dagger = (\mathbf{T}|k\rangle)^\dagger = \left(\sum_{j=1}^n \langle \bar{k} | j \rangle \langle j | \right) = \langle k | \mathbf{T}^\dagger, \text{ where: } \langle \bar{k} | j \rangle = \langle k | \mathbf{T}^\dagger | j \rangle = \langle j | \bar{k} \rangle^* \quad (2.2.5a)$$

Here Hermitian conjugate T^\dagger of an operator-matrix T is defined. True to form, its matrix representation is the transpose-conjugate of the original T -matrix. Also, from (2.2.4) it follows that the dagger-transform operator \mathbf{T}^\dagger that transforms bras is just the inverse of ket transformer \mathbf{T} .

$$\langle k | \mathbf{T}^\dagger | j \rangle = \langle j | \bar{k} \rangle^* = \langle \bar{k} | j \rangle = \langle j | \mathbf{T} | k \rangle^* = \langle k | \mathbf{T}^{-1} | j \rangle, \text{ or: } \mathbf{T}^\dagger = \mathbf{T}^{-1} \quad (2.2.5b)$$

The ability to invert a matrix by simply transposing and conjugating is a great computational luxury, especially if the matrices are large. An operator \mathbf{U} that satisfies $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ is called a *unitary operator*. Such an operator preserves the unit norm as well as all bra-ket scalar products.

For our simple example of a ϕ -rotation the bras transform as follows.

$$\left\{ \begin{aligned} \langle \bar{x} | = \langle x | \mathbf{T}^{-1} = \langle x | \cos \phi + \langle y | \sin \phi, & \quad \langle \bar{y} | = \langle y | \mathbf{T}^{-1} = -\langle x | \sin \phi + \langle y | \cos \phi \end{aligned} \right\} \quad (2.2.6a)$$

In this case the inverse is simply the transpose (\mathbf{T}^T); no conjugation is needed since the matrix is real.

$$\left(\begin{array}{cc} \langle x | \bar{x} \rangle & \langle x | \bar{y} \rangle \\ \langle y | \bar{x} \rangle & \langle y | \bar{y} \rangle \end{array} \right)^\dagger = \left(\begin{array}{cc} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{array} \right)^\dagger = \left(\begin{array}{cc} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{array} \right) = \left(\begin{array}{cc} \langle x | \mathbf{T}^\dagger | x \rangle & \langle x | \mathbf{T}^\dagger | y \rangle \\ \langle y | \mathbf{T}^\dagger | x \rangle & \langle y | \mathbf{T}^\dagger | y \rangle \end{array} \right) = \left(\begin{array}{cc} \langle \bar{x} | x \rangle & \langle \bar{x} | y \rangle \\ \langle \bar{y} | x \rangle & \langle \bar{y} | y \rangle \end{array} \right) \quad (2.2.6b)$$

(In this case, a \pm sign flip of angle ϕ would do the trick, too.) An operator \mathbf{O} that satisfies $\mathbf{O}^T = \mathbf{O}^{-1}$ is called an *orthogonal operator*. It keeps real unit vectors orthogonal as it transforms them.

(b) Bra-ket vector component transformations

Once the T-matrix is known, it is a simple matter to derive the transformation rules for components of any ket vector $|\Psi\rangle$ that lives in this ket space. It is important to remember that the abstract $|\Psi\rangle$ is not the thing that changes in a change-of-basis transformation. $|\Psi\rangle$ is just being expressed in two equivalent ways. (In other words, it has two "aliases" as shown in the equations (2.2.7) and Fig. 2.2.2 below.)

$$|\Psi\rangle = |x\rangle\langle x|\Psi\rangle + |y\rangle\langle y|\Psi\rangle = |\bar{x}\rangle\langle \bar{x}|\Psi\rangle + |\bar{y}\rangle\langle \bar{y}|\Psi\rangle \quad (2.2.7)$$

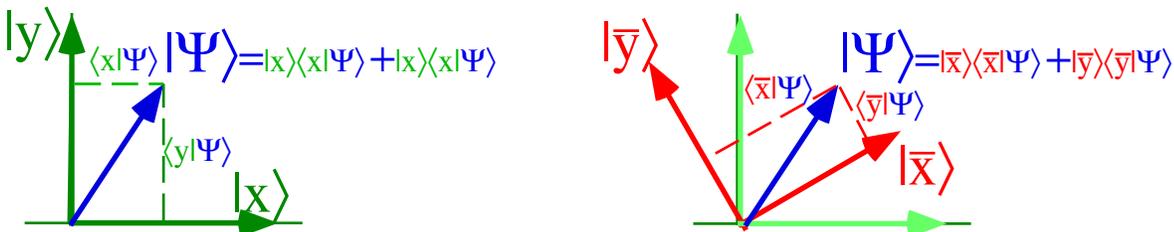


Fig. 2.2.2 Same vector $|\Psi\rangle$ with two sets of coordinate bases. ("Passive" or "Alias" transformation)

A change-of-basis transformation gives one "alias", say $(\langle \bar{x} | \Psi \rangle, \langle \bar{y} | \Psi \rangle)$, in terms of another, say $(\langle x | \Psi \rangle, \langle y | \Psi \rangle)$. Here, the transformation is obtained either by multiplying $\langle \bar{x} |$ and $\langle \bar{y} |$ in (2.2.6a) on the right by $|\Psi\rangle$ or multiplying $|\Psi\rangle$ in (2.2.7) on the left by $\langle \bar{x} |$ or $\langle \bar{y} |$. Below are the results.

$$\begin{aligned}\langle \bar{x} | \Psi \rangle &= \langle \bar{x} | x \rangle \langle x | \Psi \rangle + \langle \bar{x} | y \rangle \langle y | \Psi \rangle = \cos \phi \langle x | \Psi \rangle + \sin \phi \langle y | \Psi \rangle \\ \langle \bar{y} | \Psi \rangle &= \langle \bar{y} | x \rangle \langle x | \Psi \rangle + \langle \bar{y} | y \rangle \langle y | \Psi \rangle = -\sin \phi \langle x | \Psi \rangle + \cos \phi \langle y | \Psi \rangle\end{aligned}\quad (2.2.8a)$$

Matrix form for this is the following which uses the inverse (2.2.6b) of transformation matrix (2.2.2).

$$\begin{pmatrix} \langle \bar{x} | \Psi \rangle \\ \langle \bar{y} | \Psi \rangle \end{pmatrix} = \begin{pmatrix} \langle \bar{x} | x \rangle & \langle \bar{x} | y \rangle \\ \langle \bar{y} | x \rangle & \langle \bar{y} | y \rangle \end{pmatrix} \begin{pmatrix} \langle x | \Psi \rangle \\ \langle y | \Psi \rangle \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \langle x | \Psi \rangle \\ \langle y | \Psi \rangle \end{pmatrix}\quad (2.2.8a)$$

It is also inverse to our first change-of-basis example (1.2.6b). See (2.1.11) for general n -state formulas.

Even for a simple example like our ϕ -rotation it is difficult to derive the relations (2.2.8) just by looking at the components of a general vector in Fig. 2.2.2. (Try it!) This is one of the reasons for making the rule stated at the beginning of this Section 2.2. *You should work with bases first.*

It is important not to confuse a change-of-basis or "alias" transformation with an "active" or "alibi" transformation in which the operator \mathbf{T} is used to move a state vector $|\Psi\rangle$ into a new vector $\mathbf{T}|\Psi\rangle = |\Psi(T)\rangle$ as shown in Fig. 2.2.3. Rotation \mathbf{T} acted on the bases in Fig. 2.2.1-2 as $|\Psi\rangle$ stood still.

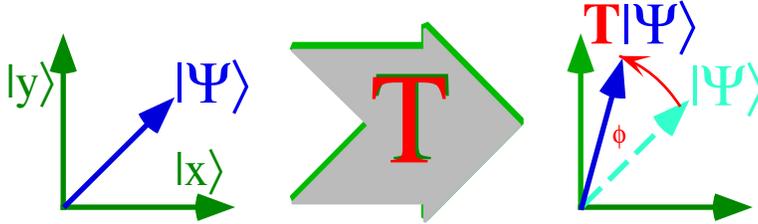


Fig. 2.2.3 Same basis but vector $|\Psi\rangle$ moves elsewhere. ("Active" or "Alibi" transformation)

Active transformation operations generally stand for analyzers or other parts of the physical space-time environment. Representations of an active transformation $\mathbf{T}|\Psi\rangle$ are made by attaching to it the bras and kets for whichever basis you want to use, say in this case, the original $\{|x\rangle, |y\rangle\}$.

$$\begin{aligned}\langle x | \mathbf{T} | \Psi \rangle &= \langle x | \mathbf{T} | x \rangle \langle x | \Psi \rangle + \langle x | \mathbf{T} | y \rangle \langle y | \Psi \rangle \\ \langle y | \mathbf{T} | \Psi \rangle &= \langle y | \mathbf{T} | x \rangle \langle x | \Psi \rangle + \langle y | \mathbf{T} | y \rangle \langle y | \Psi \rangle\end{aligned}\quad (2.2.9a)$$

The matrix form for this active transformation is

$$\begin{pmatrix} \langle x | \mathbf{T} | \Psi \rangle \\ \langle y | \mathbf{T} | \Psi \rangle \end{pmatrix} = \begin{pmatrix} \langle x | \Psi(T) \rangle \\ \langle y | \Psi(T) \rangle \end{pmatrix} = \begin{pmatrix} \langle x | \mathbf{T} | x \rangle & \langle x | \mathbf{T} | y \rangle \\ \langle y | \mathbf{T} | x \rangle & \langle y | \mathbf{T} | y \rangle \end{pmatrix} \begin{pmatrix} \langle x | \Psi \rangle \\ \langle y | \Psi \rangle \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \langle x | \Psi \rangle \\ \langle y | \Psi \rangle \end{pmatrix}\quad (2.2.9b)$$

Polarization devices that do this operation and many more will be discussed later.

(c) Group axioms

The axioms of *group theory* are mathematical axioms set down by Everiste Galois shortly before he was killed in a duel at the age of 21. They are closely related to the axioms of quantum theory which we have stated so far. Groups play a key role in the development of quantum theory, particularly at its advanced levels.

A *group* G is a set of operations $G = \{a, b, c, \dots\}$ or *elements* that can be combined in group products to give other operations in the same set. Examples are rotations or permutations, the latter of which occupied Galois' attention. Below are listed the axioms which all groups must satisfy. (The first we've already stated.)

Each is discussed as to its relevance to the set $U=\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots\}$ of all unitary transformation matrices which satisfy the quantum axioms 1-4 for an n -state system. As will be seen, U is a group which is labeled the *n -dimensional unitary group $U(n)$* .

(1) The closure axiom

Products $ab = c$ are defined between any two group elements a and b , and the result c is contained in the group.

Products $\mathbf{A}\mathbf{B}=\mathbf{C}$ of transformation operators are defined by their matrix representations as are the operators themselves according to quantum axiom-4 (The closure or completeness axiom.) as explained in Sec. 2.1b(2). In Sec. 2.2(a) it was shown that all transformation matrices are unitary. Given $\mathbf{A}^\dagger\mathbf{A} = \mathbf{1}$ and $\mathbf{B}^\dagger\mathbf{B} = \mathbf{1}$ we must prove that the product $\mathbf{A}\mathbf{B}=\mathbf{C}$ is also unitary. Inserting $\mathbf{A}^\dagger\mathbf{A} = \mathbf{1}$ between \mathbf{B}^\dagger and \mathbf{B} gives

$$\mathbf{B}^\dagger\mathbf{A}^\dagger\mathbf{A}\mathbf{B} = \mathbf{B}^\dagger\mathbf{B} = \mathbf{1} \quad (2.2.10a)$$

or

$$\mathbf{C}^\dagger\mathbf{C} = \mathbf{1} \quad (2.2.10b)$$

where showing

$$\mathbf{C}^\dagger = (\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger \quad (2.2.10c)$$

is left as an exercise.

(2) The associativity axiom

Products $(ab)c$ and $a(bc)$ are equal for all elements a , b , and c in the group .

Associativity is automatically guaranteed for matrix products on which the algebraic significance of quantum axiom-4 is based.(You should be able to prove this even if it's not a familiar axiom.)

(3) The identity axiom

There is a unique element I (the identity) such that $I\cdot a = a = a\cdot I$ for all elements a in the group ..

According to the quantum axiom-2 (the orthonormality or identity axiom) there is a unique unit transformation matrix $\langle j|k\rangle = \delta_{jk}$ that does nothing. It represents the perfect "do-nothing" analyzer. We indicate the corresponding abstract operator by $\mathbf{1}$ in the abstract completeness relation (2.1.20).

(4) The inverse axiom

For all elements a in the group there is an inverse element a^{-1} such that $a^{-1}a = I = a\cdot a^{-1}$.

This seems to follow since we already know that the Hermitian conjugate \mathbf{A}^\dagger of any operator \mathbf{A} is also its inverse $\mathbf{A}^\dagger=\mathbf{A}^{-1}$. (Recall (2.2.5b).) There is, however, a catch. For infinite state systems one cannot

guarantee that $\mathbf{A}^\dagger \mathbf{A} = \mathbf{1}$ also implies $\mathbf{A} \mathbf{A}^\dagger = \mathbf{1}$, or vice-versa. However, it does follow for finite systems. (See exercises.) In any case, for a transformation matrix satisfying quantum axioms 1-4, even with $n=\infty$, it can be proved that all operators commute with their conjugates and inverses.

This is true in spite of the fact that unitary group operators do not generally commute with each other.

$$\mathbf{AB} \neq \mathbf{BA} \text{ (for some } \mathbf{A}, \mathbf{B} \text{ in } U(n) \text{ for } n \geq 1)$$

Groups that satisfy a fifth postulate of commutivity are called *Abelian* groups.

(5) The commutative axiom (Abelian groups only)

All elements a in an Abelian group are mutually commuting: $a \cdot b = b \cdot a$.

Only the one-dimensional unitary group

$$U(1) = \{1, \dots, e^{i\alpha}, \dots\} \text{ where: } (-\pi < \alpha \leq \pi) \quad (2.2.11)$$

is Abelian. It consists of all possible phase factors for a single-state (1D) system, and these numbers obviously commute and form a group labeled by a single real parameter which is the phase angle α .

(d) U(n) group dimension

The group $U(n)$ is the set of all n -by- n transformation matrices having components $U_{ij} = T_{ij}$ each of which satisfy n^2 unitarity equations of the form

$$\sum_{j=1}^n U_{ij}^\dagger U_{jk} = (U^\dagger U)_{ik} = (\mathbf{1})_{ik} = \delta_{ik} = \sum_{j=1}^n U_{ji}^* U_{jk} \quad (2.2.12)$$

The n^2 components U_{ij} are generally complex numbers amounting to $2n^2$ real parameters. So the number of independent real parameters to label $U(n)$ operators is $2n^2 - n^2 = n^2$ and is called the unitary *group dimension*, the number of *quantum coordinates*. A 2-state system's $U(2)$ dimension is $2^2=4$.

(e) SU(n) group dimension

For most of quantum theory the over-all phase of a system is unmeasurable and ignorable. To avoid dealing with such a phase one generally restricts attention to matrices of $U(n)$ of unit determinant.

$$\det |U| = 1 \quad (2.2.13)$$

Such operators form a subgroup of $U(n)$ called the *special unimodular group* $SU(n)$. Because, of equation (2.2.13) the $SU(n)$ dimension is $n^2 - 1$, one less than that of $U(n)$. For $SU(2)$, which we study first, this dimension is $2^2 - 1 = 3$, so the number of 2-state quantum coordinates is just three. ($2^2 - 1 = 3$)

Problems for Ch. 2

Daggers

2.1.1. Prove identities for the following. Use representations to help with abstract cases.

- (a) $(\mathbf{AB})^\dagger = ?$ (In terms of $(\mathbf{A})^\dagger$ and $(\mathbf{B})^\dagger$.)
- (b) $(\mathbf{A}|\psi\rangle)^\dagger = ?$ (In terms of $(\mathbf{A})^\dagger$ and $(|\psi\rangle)^\dagger = ?$.)
- (c) $\langle\langle\phi|\psi\rangle\rangle^\dagger = ?$ (In terms of $\langle\phi|\psi\rangle^* = ?$.)
- (d) $\langle\langle\phi|\mathbf{A}|\psi\rangle\rangle^\dagger = ?$ (In terms of $(\mathbf{A})^\dagger$ and $(|\psi\rangle)^\dagger = ?$, etc..)
- (e) $(\mathbf{ABC})^\dagger = ?$ $(\mathbf{ABC})^{-1} = ?$
- (f) $\langle\langle\phi|\mathbf{ABC}|\psi\rangle\rangle^\dagger = ?$
- (g) $(|\phi\rangle\langle\psi|)^\dagger = ?$

Transforming Backwards and Forwards

2.2.1. Suppose the following basic definition of a transformation \mathbf{T} from a basis $\{|1\rangle, |2\rangle, |3\rangle\}$ to another basis $\{|1'\rangle, |2'\rangle, |3'\rangle\}$:

$$|1'\rangle = \mathbf{T}|1\rangle = (|1\rangle - |2\rangle)/\sqrt{2}, \quad |2'\rangle = \mathbf{T}|2\rangle = (|1\rangle + |2\rangle)/\sqrt{2}, \quad |3'\rangle = \mathbf{T}|3\rangle = |3\rangle/i, \quad (i = e^{\pi i/2})$$

- (a) Construct the 3×3 matrix representation of \mathbf{T} and of \mathbf{T}^\dagger in the basis $\{|1\rangle, |2\rangle, |3\rangle\}$.
- (b) Construct the 3×3 matrix representation of \mathbf{T} and of \mathbf{T}^\dagger in the basis $\{|1'\rangle, |2'\rangle, |3'\rangle\}$.
- (c) Write in matrix form a change-of-basis transformation for prime representation of a ket $|\psi\rangle$, that is, $\{\langle 1'|\psi\rangle, \langle 2'|\psi\rangle, \langle 3'|\psi\rangle\}$, in terms of its original representation $\{\langle 1|\psi\rangle, \langle 2|\psi\rangle, \langle 3|\psi\rangle\}$.
- (c)† Write in matrix form a change-of-basis transformation for prime representation of a bra $\langle\psi|$, that is, $\{\langle\psi|1'\rangle, \langle\psi|2'\rangle, \langle\psi|3'\rangle\}$, in terms of its original representation $\{\langle\psi|1\rangle, \langle\psi|2\rangle, \langle\psi|3\rangle\}$.
- (d) Write in matrix form a change-of-basis transformation for prime representation of operator \mathbf{U} , that is, $\{\langle 1'|\mathbf{U}|1'\rangle, \langle 1'|\mathbf{U}|2'\rangle, \dots\}$, in terms of its original representation $\{\langle 1|\mathbf{U}|1\rangle, \langle 1|\mathbf{U}|2\rangle, \dots\}$.
- (e) Are any of the (a-b) results..Unitary matrices? .. Hermitean matrices? ..Orthogonal matrices?
- (f) Are any of the matrices from (a) and equal to those from (b)? Which, if any, of the (a)-(b) equalities are a general result? Why or why not? (Prove or give a counter example.)

Mirror-Mirror

2.2.2. A clothing store lets you examine your new suit in a device that consists of two vertical planar mirrors. Mirror X extends along the x-axis. Mirror Φ extends along an axis that is rotated counter clockwise by angle ϕ around the vertical hinge that forms the intersection of the two mirrors at $(x,y)=(0,0)$. You stand somewhere between the two mirrors and try various ϕ while looking at any of several images of your necktie (or necklace) which is located where you're standing at $\mathbf{n}=(x,y)$. (Neglect vertical z-axis.) *Start from basic (basis-vector) definitions only to work the following questions. Deriving amplitudes directly will be marked down.*

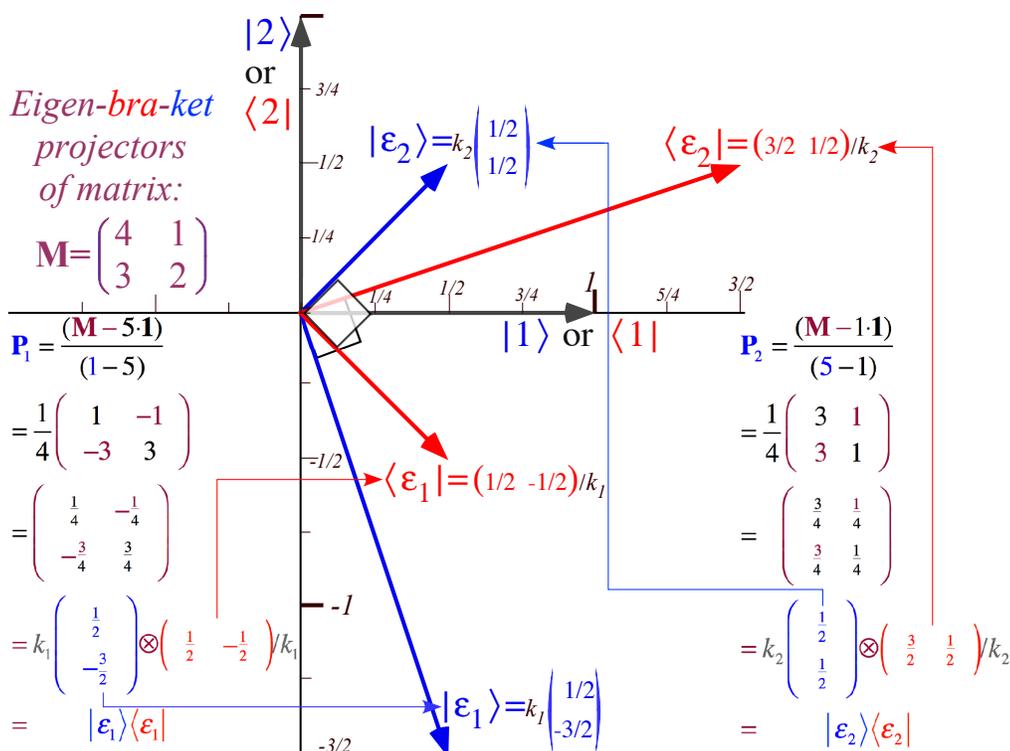
- (a) Represent the transformation $\mathbf{T}(X)$ that describes reflections by mirror X in xy-basis. Compute its effect on necktie point \mathbf{n} . Sketch top view of this mapping.
- (b) Represent the transformation $\mathbf{T}(\Phi)$ that describes reflections by mirror Φ in xy-basis. Compute its effect on necktie point \mathbf{n} . Sketch top view of this mapping.
- (c) Represent the transformation $\mathbf{T}(\Phi X)$ that describes reflections by mirror X followed by mirror Φ in the xy-basis. Compute its effect on necktie point \mathbf{n} . Sketch top view of this mapping. What familiar operation is this? Express as group product.
- (d) Represent the transformation $\mathbf{T}(X\Phi)$ that describes reflections by mirror Φ followed by mirror X in the xy-basis. Compute its effect on necktie point \mathbf{n} . Sketch top view of this mapping. What familiar operation is this? Express as group product. Do $\mathbf{T}(X)$ and $\mathbf{T}(\Phi)$ ever commute?
- (e) Calculate the determinants and trace of each of the resulting operations (a) thru (d). Comment on the physical or geometric significance, if any, of these numbers.



QM
for
AMOP
W. G. Harter

Chapter 3

Operator Eigensolutions



The concept of an operator’s “own-states” or *eigenstates* is introduced first through physical processes of analyzer filters, then visualized geometrically, and finally analyzed algebraically. The physical axioms 1-4 stated in Chapter 2 are related to four powerful *theorems* about the spectral decomposition of matrices. Applications of spectral decomposition to transformation and transfer matrices are shown.

Chapter 3 Introduction to Operator Eigensolutions

- 3.1 Operator Eigensolutions and Projection Operators3**
 - Visualizing Real Symmetric Matrices and Real Eigenvectors4
 - (a) Eigenvalue equations 5
 - (1) Secular equations 6
 - (2) Hamilton-Cayley equations 7
 - (b) Eigenvector projectors (Distinct eigenvalues) 8
 - (1) Projector normalization 9
 - Matrix products and eigensolutions for polarizer-counter arrangements11
 - (2) Projector completeness and spectral decomposition 12
 - Eigensolutions are stationary or extreme-value solutions14
 - (3) Diagonalizing transformations from projectors 15
 - Matrix products and eigensolutions for active analyzers16
 - (c) Eigenvector projectors (Degenerate eigenvalues) 17
 - (1) Minimal equation and diagonalizability criterion 17
 - (2) Nilpotent operators ("Bad" degeneracy) 18
 - (3) Multiple diagonalization ("Good" degeneracy) 19
 - Gram-Schmidt orthogonalization20
 - (d) Projector splitting: A key to algebraic reduction 20
 - (e) Why symmetry groups are useful 23
 - Quadratic surfaces help to visualize matrix operations24
- 3.2 Approximate Eigensolutions by Perturbation Techniques25**
 - (a) Secular determinantal expansion 25
 - (b) Perturbation approximations 26
 - (c) Testing perturbation approximation with exact 2x2 eigenvalues 28
 - Appendix 3.A Matrix Determinants, Adjuncts, and Inverses 1
 - Appendix 3.B Classification of Permutations 4
- UNIT. 1 REVIEW TOPICS AND FORMULAS1**

Problems for Chapter 3.

3.1 Operator Eigensolutions and Projection Operators

Many quantum processes and analyzers can be represented by complex matrix *transfer* operators \mathbf{T} that act on an input state ket $|\Psi_{IN}\rangle$ to give the resulting output state kets $|\Psi_{OUT}\rangle$ as follows.

$$|\Psi_{OUT}\rangle = \mathbf{T}|\Psi_{IN}\rangle \tag{3.1.1}$$

(Recall discussion around (1.3.2).) In this way, matrix products predict the effect of the corresponding T-analyzer or a whole chain of analyzers. Generally, the effect of an analyzer is to change a state $|\Psi\rangle$ to one whose output vector $\mathbf{T}|\Psi\rangle$ is rotated or otherwise transformed as shown in Fig. 3.1.1 below. A transfer operator \mathbf{T} that is unitary ($\mathbf{T}^\dagger = \mathbf{T}^{-1}$) is also a *transformation* operator and satisfies Axioms 1-4.

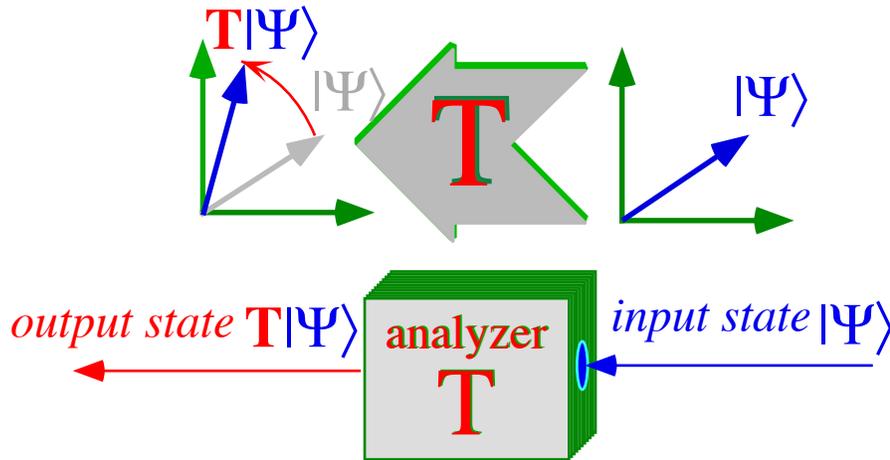


Fig. 3.1.1 Effect of analyzer represented by ket vector transformation of $|\Psi\rangle$ to new vector $\mathbf{T}|\Psi\rangle$.

However, most analyzers have certain of their *own* states whose kets $|e_j\rangle$ lie along certain "magic" directions that do not change when \mathbf{T} acts on them, that is, the input ket $|e_j\rangle$ just gets multiplied by a phase factor or other number ϵ_j as in

$$\mathbf{T}|e_j\rangle = \epsilon_j |e_j\rangle, \tag{3.1.2}$$

but the vector $|e_j\rangle$ remains pointing in the same direction as shown in Fig. 3.1.2 below.

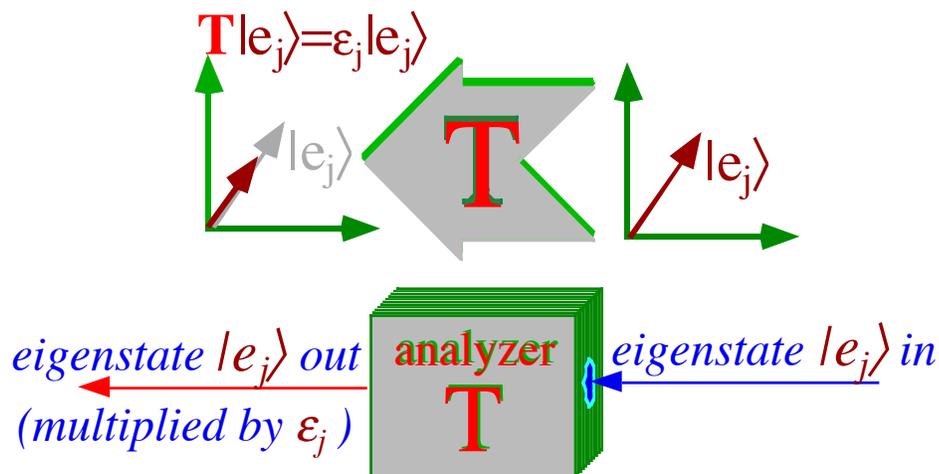
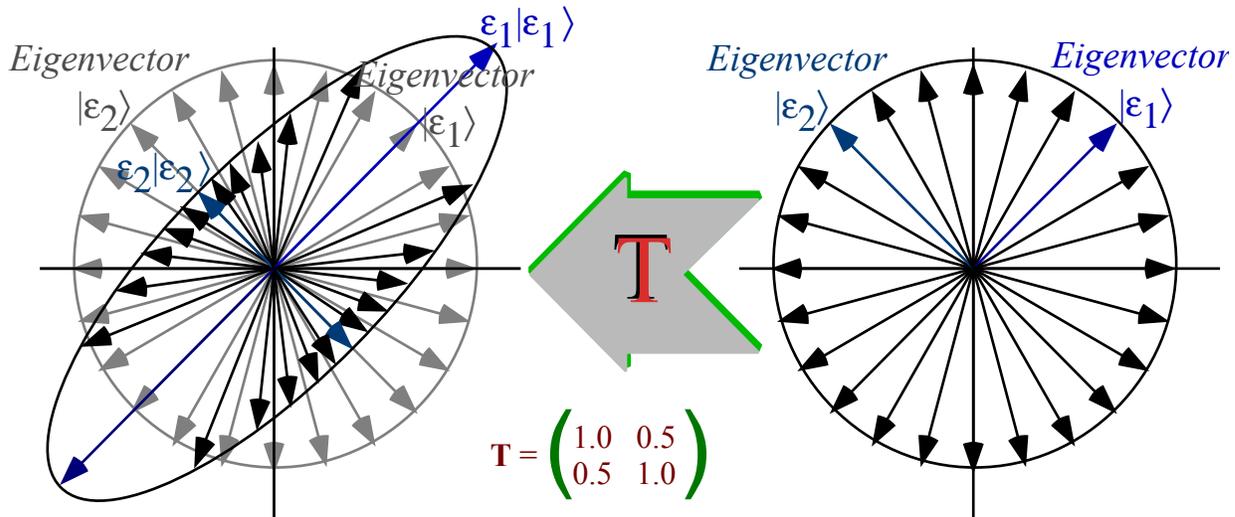


Fig. 3.1.2 Effect of analyzer on eigenket $|e_j\rangle$ is to simply multiply by eigenvalue ϵ_j ($\mathbf{T}|e_j\rangle = \epsilon_j |e_j\rangle$).

Visualizing Real Symmetric Matrices and Real Eigenvectors

You can learn something about a real matrix operator or transformation \mathbf{T} by applying it to a circular array of unit vectors \mathbf{c} . As shown below a matrix $\mathbf{T} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$ maps a circular array into an elliptical one



Only two vectors in the upper half plane survive the transformation \mathbf{T} without changing their directions. These lucky vectors are the *eigenvectors*

$$|\epsilon_1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix} / \sqrt{2}, \quad \text{and} \quad |\epsilon_2\rangle = \begin{pmatrix} -1 \\ 1 \end{pmatrix} / \sqrt{2}$$

which transform as follows: $\mathbf{T}|\epsilon_1\rangle = \epsilon_1|\epsilon_1\rangle = 1.5|\epsilon_1\rangle$, and $\mathbf{T}|\epsilon_2\rangle = \epsilon_2|\epsilon_2\rangle = 0.5|\epsilon_2\rangle$ by only suffering a length change given by *eigenvalues* $\epsilon_1 = 1.5$ and $\epsilon_2 = 0.5$, respectively. Obviously, the negatives $-|\epsilon_1\rangle$ or $-|\epsilon_2\rangle$ of eigenvectors are eigenvectors, too, as is $17|\epsilon_1\rangle$ or $-29|\epsilon_2\rangle, \dots$ etc. *Normalization* ($\langle \mathbf{c} | \mathbf{c} \rangle = 1$) is a separate condition that we generally require of eigenvectors, too.

Each vector $|\mathbf{r}\rangle$ on the left hand ellipse maps back to a vector $|\mathbf{c}\rangle = \mathbf{T}^{-1}|\mathbf{r}\rangle$ on the right hand unit circle, Each $|\mathbf{c}\rangle$ has unit length: $\langle \mathbf{c} | \mathbf{c} \rangle = 1 = \langle \mathbf{r} | \mathbf{T}^{-1} \mathbf{T}^{-1} | \mathbf{r} \rangle = \langle \mathbf{r} | \mathbf{T}^{-2} | \mathbf{r} \rangle$. (\mathbf{T} is real-symmetric: $\mathbf{T}^\dagger = \mathbf{T} = \mathbf{T}^T$.)

$$\mathbf{c} \cdot \mathbf{c} = 1 = \mathbf{r} \cdot \mathbf{T}^{-2} \cdot \mathbf{r} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} T_{xx} & T_{xy} \\ T_{yx} & T_{yy} \end{pmatrix}^{-2} \begin{pmatrix} x \\ y \end{pmatrix}$$

This simplifies if rewritten in a coordinate system (x_1, x_2) of eigenvectors $|\epsilon_1\rangle$ and $|\epsilon_2\rangle$ where $\mathbf{T}^{-2}|\epsilon_1\rangle = \epsilon_1^{-2}|\epsilon_1\rangle$ and $\mathbf{T}^{-2}|\epsilon_2\rangle = \epsilon_2^{-2}|\epsilon_2\rangle$, that is, \mathbf{T} , \mathbf{T}^{-1} , and \mathbf{T}^{-2} are each represented by a diagonal matrix.

$$\begin{pmatrix} \langle \epsilon_1 | \mathbf{T} | \epsilon_1 \rangle & \langle \epsilon_1 | \mathbf{T} | \epsilon_2 \rangle \\ \langle \epsilon_2 | \mathbf{T} | \epsilon_1 \rangle & \langle \epsilon_2 | \mathbf{T} | \epsilon_2 \rangle \end{pmatrix} = \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \langle \epsilon_1 | \mathbf{T}^{-2} | \epsilon_1 \rangle & \langle \epsilon_1 | \mathbf{T}^{-2} | \epsilon_2 \rangle \\ \langle \epsilon_2 | \mathbf{T}^{-2} | \epsilon_1 \rangle & \langle \epsilon_2 | \mathbf{T}^{-2} | \epsilon_2 \rangle \end{pmatrix} = \begin{pmatrix} \epsilon_1^{-2} & 0 \\ 0 & \epsilon_2^{-2} \end{pmatrix}$$

So, the matrix equation simplifies to an elementary ellipse equation of the form $(x/a)^2 + (y/b)^2 = 1$.

$$\mathbf{c} \cdot \mathbf{c} = 1 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \epsilon_1^{-2} & 0 \\ 0 & \epsilon_2^{-2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left(\frac{x_1}{\epsilon_1}\right)^2 + \left(\frac{x_2}{\epsilon_2}\right)^2$$

The ellipse semi-major-minor axes are eigenvalues $\epsilon_1 = 1.5$ and $\epsilon_2 = 0.5$. The axes are tilted as shown above. Such a \mathbf{T} operation is a *tensor* operation. \mathbf{T} anisotropically stretches and squeezes the space.

(a) Eigenvalue equations

The equation (3.1.2) is called an *eigenvalue equation* $\mathbf{T} |\epsilon_j\rangle = \epsilon_j |\epsilon_j\rangle$ and the kets $|\epsilon_j\rangle$ are called ket eigenvectors or simply *eigenkets* $|\epsilon_j\rangle$ and the scalar numbers are called *eigenvalues* ϵ_j of operator \mathbf{T} . For analyzers described by unitary operators ($\mathbf{T}^\dagger = \mathbf{T}^{-1}$) the eigenvalues are simply phase factors

$$\epsilon_j = e^{i\phi}, \tag{3.1.3}$$

so eigenkets stay the same magnitude. If the analyzer has a counter or particle sources then it may decrease (as in the Fig. 3.1.2) or increase the magnitude (and probability) of an eigenket vector.

The prefix "eigen" means "own" in German. The eigenvectors of a single analyzer-T are its own vectors, literally. We would call them "ownvectors" if we had to purge German from English. The eigenvectors $\{ |\epsilon_1\rangle, |\epsilon_2\rangle, \dots \}$ correspond to the *eigenstates* that get sorted out inside an analyzer as in Fig. 1.3.1 or Fig. 1.3.8. If a T-analyzer is set to a filter configuration like Fig. 1.3.4 then it can produce a beam that is made purely of one or another of its own eigenstates $|\epsilon_j\rangle$. (Excuse the bilingual redundancy.) Then another T-analyzer in the "do-nothing" mode would pass each of the resulting $|\epsilon_j\rangle$ particles 100% unchanged (except maybe for an overall phase) according to axiom-3. That would be an example of an eigen-equation $\mathbf{T} |\epsilon_j\rangle = \epsilon_j |\epsilon_j\rangle$ in its purest form.

If the \mathbf{T} -operator is represented in its own eigenbasis (Sorry, another bilingual redundancy.) then its matrix representation takes has a very simple *diagonal* form according to axiom-3.

$$\langle \epsilon_i | \mathbf{T} | \epsilon_j \rangle = \epsilon_j \langle \epsilon_i | \epsilon_j \rangle = \epsilon_j \delta_{ij} \tag{3.1.4a}$$

The diagonal matrix for an n-state system is

$$\begin{pmatrix} \langle \epsilon_1 | \mathbf{T} | \epsilon_1 \rangle & \langle \epsilon_1 | \mathbf{T} | \epsilon_2 \rangle & \dots & \langle \epsilon_1 | \mathbf{T} | \epsilon_n \rangle \\ \langle \epsilon_2 | \mathbf{T} | \epsilon_1 \rangle & \langle \epsilon_2 | \mathbf{T} | \epsilon_2 \rangle & \dots & \langle \epsilon_2 | \mathbf{T} | \epsilon_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \epsilon_n | \mathbf{T} | \epsilon_1 \rangle & \langle \epsilon_n | \mathbf{T} | \epsilon_2 \rangle & \dots & \langle \epsilon_n | \mathbf{T} | \epsilon_n \rangle \end{pmatrix} = \begin{pmatrix} \epsilon_1 & 0 & \dots & 0 \\ 0 & \epsilon_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \epsilon_n \end{pmatrix} \tag{3.1.4b}$$

However, we are usually given the \mathbf{T} -operator in someone else's basis $\{ |1\rangle, |2\rangle, \dots, |n\rangle \}$ as in

$$\begin{pmatrix} \langle 1 | \mathbf{T} | 1 \rangle & \langle 1 | \mathbf{T} | 2 \rangle & \dots & \langle 1 | \mathbf{T} | n \rangle \\ \langle 2 | \mathbf{T} | 1 \rangle & \langle 2 | \mathbf{T} | 2 \rangle & \dots & \langle 2 | \mathbf{T} | n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N | \mathbf{T} | 1 \rangle & \langle N | \mathbf{T} | 2 \rangle & \dots & \langle N | \mathbf{T} | n \rangle \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & \dots & T_{1n} \\ T_{21} & T_{22} & \dots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \dots & T_{nn} \end{pmatrix} \tag{3.1.4c}$$

Then the problem is one of *diagonalization* which consists of using n^2 matrix T_{ij} numbers to solve the following problems:

(Problem A) Find \mathbf{T} 's eigenvalues $\{ \epsilon_1, \epsilon_2, \dots, \epsilon_n \}$ (Find n numbers ϵ_j)

(Problem B) Find \mathbf{T} 's eigenket basis $\{ |\epsilon_1\rangle, |\epsilon_2\rangle, \dots, |\epsilon_n\rangle \}$ (Find n^2 numbers $\langle i | \epsilon_j \rangle$)

The lions share of work and information (particularly for large n) is in the n^2 components $\langle i | \epsilon_j \rangle$ of the *diagonalization transformation (d-tran) matrix* which will reduce (3.1.4c) to diagonal form (3.1.4b). The d-tran matrix $\langle i | \epsilon_j \rangle$ has in its columns the desired eigenkets $\{ |\epsilon_1\rangle, |\epsilon_2\rangle, \dots, |\epsilon_n\rangle \}$. (Recall (2.1.8).)

(1) Secular equations

The eigenvalue equations (3.1.2) for a general matrix operator \mathbf{M} can be written as follows

$$\mathbf{M}|\varepsilon_k\rangle = \varepsilon_k|\varepsilon_k\rangle, \text{ or: } (\mathbf{M} - \varepsilon_k\mathbf{1})|\varepsilon_k\rangle = \mathbf{0} \quad (3.1.5a)$$

and represented by

$$\begin{pmatrix} \langle 1|\mathbf{M}|1\rangle & \langle 1|\mathbf{M}|2\rangle & \cdots & \langle 1|\mathbf{M}|N\rangle \\ \langle 2|\mathbf{M}|1\rangle & \langle 2|\mathbf{M}|2\rangle & \cdots & \langle 2|\mathbf{M}|N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|\mathbf{M}|1\rangle & \langle N|\mathbf{M}|2\rangle & \cdots & \langle N|\mathbf{M}|N\rangle \end{pmatrix} \begin{pmatrix} \langle 1|\varepsilon_j\rangle \\ \langle 2|\varepsilon_j\rangle \\ \vdots \\ \langle N|\varepsilon_j\rangle \end{pmatrix} = \varepsilon_j \begin{pmatrix} \langle 1|\varepsilon_j\rangle \\ \langle 2|\varepsilon_j\rangle \\ \vdots \\ \langle N|\varepsilon_j\rangle \end{pmatrix} \quad (3.1.5b)$$

or

$$\begin{pmatrix} \langle 1|\mathbf{M}|1\rangle - \varepsilon_j & \langle 1|\mathbf{M}|2\rangle & \cdots & \langle 1|\mathbf{M}|N\rangle \\ \langle 2|\mathbf{M}|1\rangle & \langle 2|\mathbf{M}|2\rangle - \varepsilon_j & \cdots & \langle 2|\mathbf{M}|N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|\mathbf{M}|1\rangle & \langle N|\mathbf{M}|2\rangle & \cdots & \langle N|\mathbf{M}|N\rangle - \varepsilon_j \end{pmatrix} \begin{pmatrix} \langle 1|\varepsilon_j\rangle \\ \langle 2|\varepsilon_j\rangle \\ \vdots \\ \langle N|\varepsilon_j\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (3.1.5c)$$

These amount to n equations for each of n eigenvalues $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ or n^2 equations in all. The eigenvalues may be found by demanding that the determinant of the matrix in (3.1.5c) be zero. This is called the *secular equation*

$$\det|\mathbf{M} - \varepsilon\mathbf{1}| = 0 = (-1)^n (\varepsilon^n + a_1\varepsilon^{n-1} + a_2\varepsilon^{n-2} + \dots + a_{n-1}\varepsilon + a_n) \quad (3.1.5d)$$

where the polynomial coefficients are

$$a_1 = -\text{Trace}\mathbf{M}, \dots, a_k = (-1)^k \sum \text{diagonal k-by-k minors of } \mathbf{M}, \dots, a_n = (-1)^n \det|\mathbf{M}| \quad (3.1.5e)$$

The secular equation has n -factors, one for each eigenvalue.

$$\det|\mathbf{M} - \varepsilon\mathbf{1}| = 0 = (-1)^n (\varepsilon - \varepsilon_1)(\varepsilon - \varepsilon_2)\cdots(\varepsilon - \varepsilon_n) \quad (3.1.5f)$$

It may help to see some examples. For a two-by-two matrix \mathbf{H} ($n=2$) diagonalization is comparatively simple. The ($n=2$) secular equation is:

$$0 = \det \begin{vmatrix} H_{11} - \varepsilon & H_{12} \\ H_{21} & H_{22} - \varepsilon \end{vmatrix} = \varepsilon^2 - (H_{11} + H_{22})\varepsilon + (H_{11}H_{22} - H_{12}H_{21}) \quad (3.1.5)_{\text{example}}$$

and the polynomial coefficients are just related to matrix trace and determinant.

$$\begin{aligned} a_1 &= -(H_{11} + H_{22}) = -\text{Trace}H \\ a_2 &= H_{11}H_{22} - H_{12}H_{21} = \det H \end{aligned} \quad (3.1.5e)_{\text{example}}$$

Had we done this with a diagonal matrix then the coefficients in terms of eigenvalues would be

$$\begin{aligned} a_1 &= -(H_{11} + H_{22}) = -(\varepsilon_1 + \varepsilon_2) \\ a_2 &= H_{11}H_{22} - 0 = \varepsilon_1\varepsilon_2 \end{aligned} \quad (3.1.5e)_{\text{example}}$$

The preceding two equations must give the same numbers because the secular equation and its roots must not depend on the basis used to represent the abstract operator. Trace, determinant, and a_j are *invariant*.

For numerical examples, let us use two different matrices given below. One is Hermitian (self-conjugate) and one is not. (You might call them "good-cop" and "bad-cop", respectively.)

$$\mathbf{H} = \begin{pmatrix} 4 & -i\sqrt{3} \\ i\sqrt{3} & 2 \end{pmatrix} = \mathbf{H}^\dagger \quad (3.1.6a)$$

$$\mathbf{K} = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \neq \mathbf{K}^\dagger \quad (3.1.6b)$$

They both have the same secular equation:

$$\begin{aligned} 0 &= S(\varepsilon) = \varepsilon^2 - 6\varepsilon + 5 \\ 0 &= (\varepsilon - 1)(\varepsilon - 5) \end{aligned} \quad (3.1.7)$$

and same roots or eigenvalues ($\varepsilon_1 = 1$ and $\varepsilon_2 = 5$). However, the "bad-cop" matrix \mathbf{K} is not one that you are likely to see in quantum theory since it is neither unitary nor Hermitian. Still, it is instructive to see what the diagonalization formalism does with a pathological case such as this one. The "good-cop" matrix is not unitary, so it won't represent ideal analyzers, but because it is Hermitian, it could show up in other roles such as density operator or Hamiltonian matrices. (To be discussed later)

(2) Hamilton-Cayley equations

If each variable ε in the secular equation (3.1.5f) is replaced by the matrix operator \mathbf{M} and each ε_k by $\varepsilon_k \mathbf{1}$ then the following matrix equation results.

$$\mathbf{0} = (\mathbf{M} - \varepsilon_1 \mathbf{1})(\mathbf{M} - \varepsilon_2 \mathbf{1}) \cdots (\mathbf{M} - \varepsilon_n \mathbf{1}) \quad (3.1.8)$$

This operator equation is known as the *Hamilton-Cayley (HC) equation* or *Hamilton-Cayley theorem*.

The HC-equation is obviously true if \mathbf{M} has the diagonal form of (3.1.4b). But, that is circular logic since one needs to *prove* the diagonal form is possible first. We shall arrive at this proof in a roundabout way. For now a quick check of the HC-equation for the "bad-cop" \mathbf{K} -matrix (3.1.6b) is done below.

$$\begin{aligned} \mathbf{K}^2 - 6\mathbf{K} + 5\mathbf{1} &= \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}^2 - 6 \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} + 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 19 & 6 \\ 18 & 7 \end{pmatrix} - \begin{pmatrix} 24 & 6 \\ 18 & 12 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \left[\begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix} \right] \left[\begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix} \right] \end{aligned} \quad (3.1.8)_{\text{example}}$$

The HC-equation works fine in this case, as it does for all matrices.

(b) Eigenvector projectors (Distinct eigenvalues)

To obtain eigenvectors we construct *projection operators* \mathbf{p}_k by replacing k -th factor $(\mathbf{M} - \epsilon_k \mathbf{1})$ from HC eq.(3.1.8) by unit matrix $(\mathbf{1})$ as follows. (We assume distinct eigenvalues $\epsilon_1 \neq \epsilon_2 \neq \dots$ here.)

$$\begin{aligned} \mathbf{p}_1 &= (\mathbf{1}) (\mathbf{M} - \epsilon_2 \mathbf{1}) \cdots (\mathbf{M} - \epsilon_n \mathbf{1}) \\ \mathbf{p}_2 &= (\mathbf{M} - \epsilon_1 \mathbf{1}) (\mathbf{1}) \cdots (\mathbf{M} - \epsilon_n \mathbf{1}) \quad \text{or: } \mathbf{p}_k = \prod_{j \neq k} (\mathbf{M} - \epsilon_j \mathbf{1}) \\ &\vdots \\ \mathbf{p}_n &= (\mathbf{M} - \epsilon_1 \mathbf{1}) (\mathbf{M} - \epsilon_2 \mathbf{1}) \cdots (\mathbf{1}) \end{aligned} \tag{3.1.9}$$

Each operator \mathbf{p}_k has a delightful property. The \mathbf{p}_k solve the original eigenvector equation (3.1.5a).

$$(\mathbf{M} - \epsilon_k \mathbf{1}) \mathbf{p}_k = \mathbf{0} \quad \text{or:} \quad \mathbf{M} \mathbf{p}_k = \epsilon_k \mathbf{p}_k \tag{3.1.10a}$$

$$\mathbf{p}_k (\mathbf{M} - \epsilon_k \mathbf{1}) = \mathbf{0} \quad \text{or:} \quad \mathbf{p}_k \mathbf{M} = \epsilon_k \mathbf{p}_k \tag{3.1.10b}$$

This is true because putting back the k -th factor $(\mathbf{M} - \epsilon_k \mathbf{1})$ restores the original HC-equation and gives zero. Relation $\mathbf{M} \mathbf{p}_k = \epsilon_k \mathbf{p}_k$ implies that \mathbf{p}_k contains ket eigenvectors $|\epsilon_j\rangle$ in its columns and $\mathbf{p}_k \mathbf{M} = \epsilon_k \mathbf{p}_k$ implies that bra eigenvectors $\langle \epsilon_j|$ in its rows. (The "soft-bra-ket" notation $(|$ or $|)$ denotes un-normalized left or right eigenvectors.) Consider the "bad cop" example again. First, here are its projectors worked out.

$$\mathbf{K} - 5 \cdot \mathbf{1} = \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix} = \mathbf{p}_1, \quad \mathbf{K} - 1 \cdot \mathbf{1} = \left[\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix} = \mathbf{p}_2. \tag{3.1.9} \text{example}$$

Note that matrix \mathbf{K} eigenvector relations are satisfied many ways by the \mathbf{p}_j 's. Here are the "right handed" un-normalized $|\epsilon_j\rangle$ -ket solutions.

$$\begin{aligned} \mathbf{K} \cdot \mathbf{p}_1 &= 1 \cdot \mathbf{p}_1, & \mathbf{K} \cdot \mathbf{p}_2 &= 5 \cdot \mathbf{p}_2, \\ \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix} &= 1 \cdot \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix}, & \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix} &= 5 \cdot \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix} \\ \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} &= 1 \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix}, & \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \end{pmatrix} &= 5 \cdot \begin{pmatrix} 3 \\ 3 \end{pmatrix} \end{aligned} \tag{3.1.10a} \text{example}$$

Here are the "left handed" or $\langle \epsilon_j|$ -bra solutions.

$$\begin{aligned} \mathbf{p}_1 \cdot \mathbf{K} &= 1 \cdot \mathbf{p}_1, & \mathbf{p}_2 \cdot \mathbf{K} &= 5 \cdot \mathbf{p}_2, \\ \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} &= 1 \cdot \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix}, & \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} &= 5 \cdot \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix} \\ \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} &= 1 \cdot \begin{pmatrix} -1 & 1 \end{pmatrix}, & \begin{pmatrix} 3 & 1 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} &= 5 \cdot \begin{pmatrix} 3 & 1 \end{pmatrix} \end{aligned} \tag{3.1.10b} \text{example}$$

This is a powerful way to calculate eigenbras and eigenkets. But, there is much more power hidden in this approach. We're just getting started! Read on.

(1) Projector normalization

We may normalize \mathbf{p}_k operators to make the *idempotent projection operators* \mathbf{P}_k defined by

$$\mathbf{P}_k = \frac{\prod_{j \neq k} (\mathbf{M} - \varepsilon_j \mathbf{1})}{\prod_{j \neq k} (\varepsilon_k - \varepsilon_j)} \tag{3.1.11}$$

The *normalized* projectors are denoted by upper case \mathbf{P} and satisfy \mathbf{p} -eigen-equations (3.1.10), too.

$$\mathbf{M} \mathbf{P}_k = \varepsilon_k \mathbf{P}_k, \quad (\mathbf{M})^2 \mathbf{P}_k = (\varepsilon_k)^2 \mathbf{P}_k \dots \tag{3.1.12a}$$

$$\mathbf{P}_k \mathbf{M} = \varepsilon_k \mathbf{P}_k, \quad \mathbf{P}_k (\mathbf{M})^2 = (\varepsilon_k)^2 \mathbf{P}_k \dots \text{ etc.} \tag{3.1.12b}$$

This normalization make \mathbf{P} 's *idempotent* ($\mathbf{P}^2 = \mathbf{P}$) as follows using: $f(\mathbf{M})\mathbf{P}_k = f(\varepsilon_k)\mathbf{P}_k$.

$$\mathbf{P}_k \mathbf{P}_k = \frac{\prod_{j \neq k} (\mathbf{M} - \varepsilon_j \mathbf{1})}{\prod_{j \neq k} (\varepsilon_k - \varepsilon_j)} \mathbf{P}_k = \frac{\prod_{j \neq k} (\varepsilon_k - \varepsilon_j)}{\prod_{j \neq k} (\varepsilon_k - \varepsilon_j)} \mathbf{P}_k = \mathbf{P}_k \tag{3.1.13a}$$

A *projector orthonormalization relation* follows since projectors, normalized or not, are mutually orthogonal.

$$\mathbf{p}_j \mathbf{p}_k = \mathbf{0} \text{ for } j \neq k \quad \text{or:} \quad \mathbf{P}_j \mathbf{P}_k = \delta_{jk} \mathbf{P}_k \tag{3.3.13b}$$

Consider a "bad cop" example after (3.1.10) to begin seeing what power \mathbf{P}_k -normalization relations give.

$$\mathbf{P}_1 = \frac{\begin{pmatrix} -1 & 1 \\ 3 & -3 \\ 1 & -5 \end{pmatrix}}{(1-5)} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} \quad \mathbf{P}_2 = \frac{\begin{pmatrix} 3 & 1 \\ 3 & 1 \\ 5 & -1 \end{pmatrix}}{(5-1)} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \tag{3.1.11}\text{example}$$

Idempotence implies 4 ready-made sets of scalar products between 2 rows (bras) and 2 columns (kets).

$$\mathbf{P}_1 = \mathbf{P}_1 \mathbf{P}_1 = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix}, \quad \mathbf{P}_2 = \mathbf{P}_2 \mathbf{P}_2 = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \tag{3.1.13}\text{examples}$$

For example, (1st row)·(1st column) dot product $(\varepsilon_k | \varepsilon_k)$ is already given by \mathbf{P}_k -matrix element $(\mathbf{P}_k)_{11}$.

$$(\varepsilon_1 | \varepsilon_1) = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ -\frac{3}{4} \end{pmatrix} = \frac{1}{4} = (\mathbf{P}_1)_{11}, \quad (\varepsilon_2 | \varepsilon_2) = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \end{pmatrix} = \frac{3}{4} = (\mathbf{P}_2)_{11}.$$

One way to get a *normalized bra* $\langle \varepsilon_k |$ or *ket* $|\varepsilon_k\rangle$ is to divide un-normal row- $(\varepsilon_k |$ or column- $|\varepsilon_j)$ by $\sqrt{(\mathbf{P}_k)_{11}}$.

$$|\varepsilon_1\rangle = \frac{|\varepsilon_1\rangle}{\sqrt{(\mathbf{P}_1)_{11}}} = \frac{1}{\sqrt{\frac{1}{4}}} \begin{pmatrix} \frac{1}{4} \\ -\frac{3}{4} \end{pmatrix}, \quad \langle \varepsilon_1| = \frac{\langle \varepsilon_1|}{\sqrt{(\mathbf{P}_1)_{11}}} = \frac{\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \end{pmatrix}}{\sqrt{\frac{1}{4}}}, \quad |\varepsilon_2\rangle = \frac{|\varepsilon_2\rangle}{\sqrt{(\mathbf{P}_2)_{11}}} = \frac{1}{\sqrt{\frac{3}{4}}} \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \end{pmatrix}, \quad \langle \varepsilon_2| = \frac{\langle \varepsilon_2|}{\sqrt{(\mathbf{P}_2)_{11}}} = \frac{\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}}{\sqrt{\frac{3}{4}}}.$$

A more elegant way is to \otimes -factor (Recall (2.1.22)) each \mathbf{P}_j -matrix as shown here or in Fig. 3.1.3 below.

$$\mathbf{P}_1 = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} = k_1 \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{2} \end{pmatrix} \otimes \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} / k_1 \tag{3.1.14a} \quad \mathbf{P}_2 = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} = k_2 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \otimes \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \end{pmatrix} / k_2 \tag{3.1.14b}$$

$$= |\varepsilon_1\rangle \langle \varepsilon_1| \quad = |\varepsilon_2\rangle \langle \varepsilon_2|$$

This way shows you may shift magnitude and phase between a ket-factor $|\varepsilon_j\rangle$ and its companion bra-factor $\langle \varepsilon_j|$ by varying a *gauge adjustment constant* k_j to any non-zero value, real or complex. All the projection algebra discussed below is invariant to k_j . (The first way above has $(k_1=1, k_2=1/\sqrt{3})$. Figure 3.1.3 has $k_1=1=k_2$.)

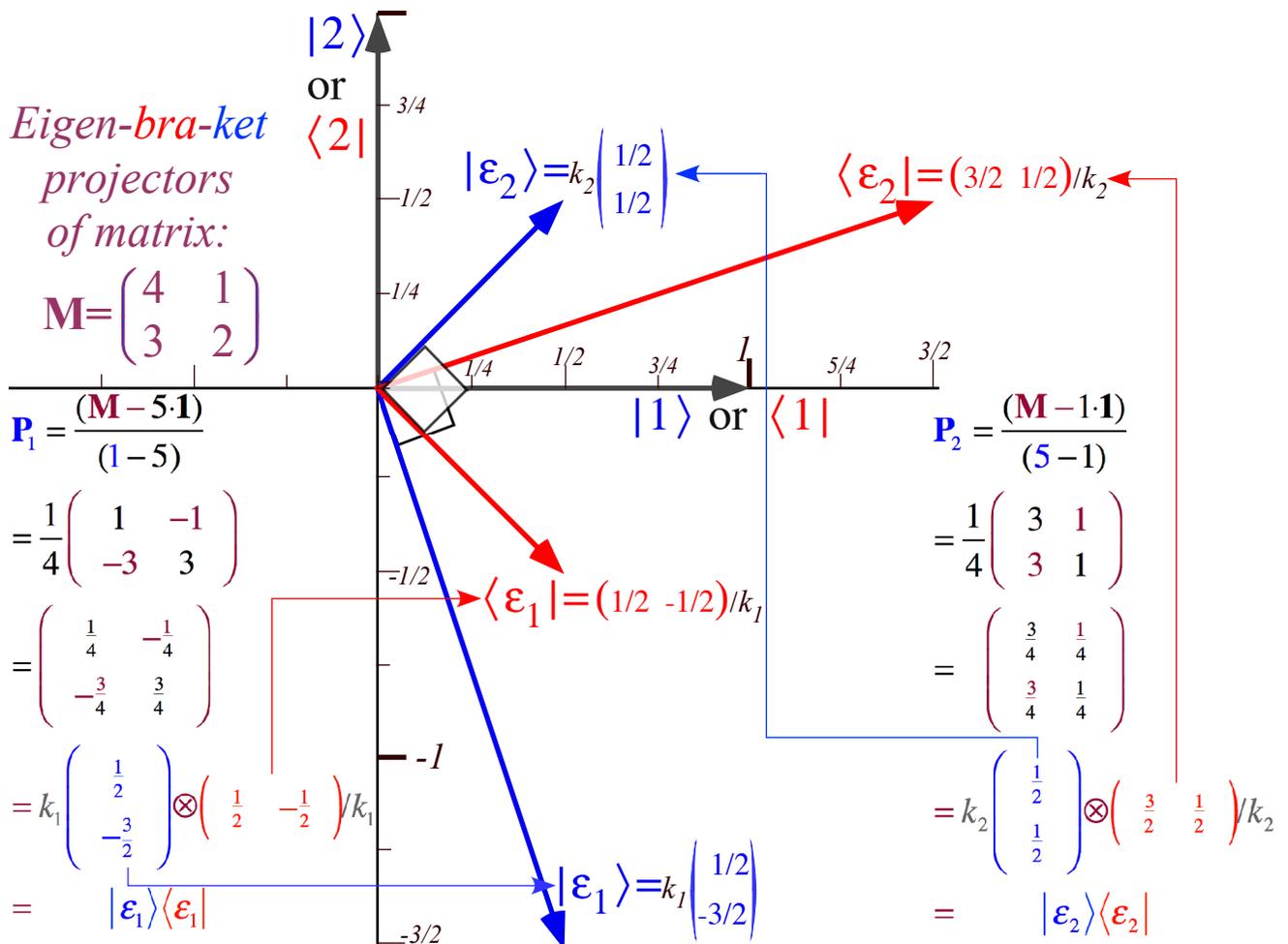


Fig. 3.1.3 Normalized eigenbras and eigenkets for asymmetric "bad cop" matrix K. (Here $k_1=1=k_2$.)

The first thing to notice is that the "bad cop" eigenvectors do not satisfy quantum conjugation relations associated with axiom-2, that is eigenbra $\langle \epsilon_j|$ is NOT equal to Hermitian conjugate $|\epsilon_j\rangle^\dagger$ of eigenket $|\epsilon_j\rangle$. Still, they are orthonormal and satisfy axiom-3. (They satisfy the "letter of the law" but not the intent, just like a "bad cop" would!) Each $\langle \epsilon_j|$ is 90° from $|\epsilon_{k \neq j}\rangle$ in Fig. 3.1.3 and normalized as per (3.1.13b) regardless of k_j .

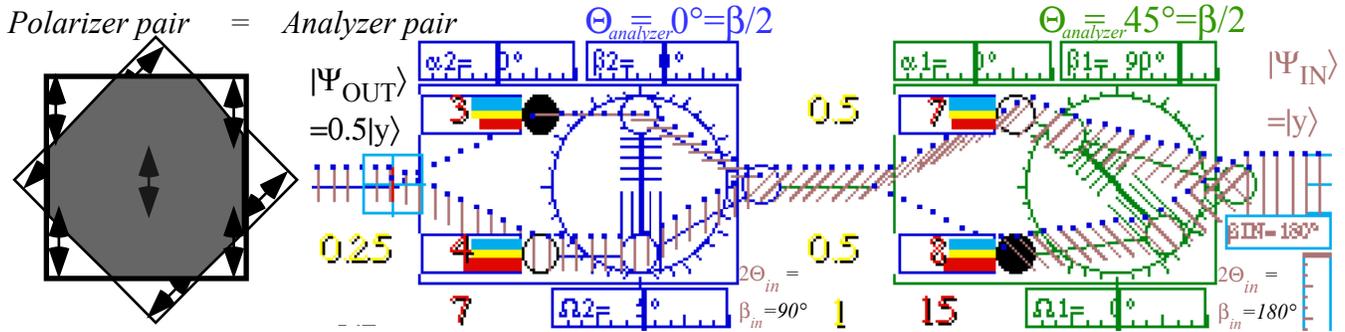
$$\langle \epsilon_1 | \epsilon_2 \rangle = 0 = \langle \epsilon_2 | \epsilon_1 \rangle \qquad \langle \epsilon_1 | \epsilon_1 \rangle = 1 = \langle \epsilon_2 | \epsilon_2 \rangle .$$

If you want to double the length of the first ket $|\epsilon_1\rangle$ in Fig. 3.1.3, you may increase k_1 from 1 to 2. Note this will halve the length of the first bra $\langle \epsilon_1|$ so the scalar products and projectors do not change. It is remarkable that you may then vary the second ket $|\epsilon_2\rangle$ and bra $\langle \epsilon_2|$ similarly and independently of the first. These are simple examples of *gauge transformations* that revise length or wave amplitude standards for plots like Fig. 3.1.3.

You should calculate **P**-matrices for the complex ("good cop") matrix **H** in (3.1.6a), and show that their eigenbra-kets also satisfy both the full quantum orthonormality and conjugation relations. The correct name for a "good cop" matrix is a *normal* matrix **N**, which simply means it commutes with its \dagger -conjugate: $\mathbf{N}^\dagger \mathbf{N} = \mathbf{N} \mathbf{N}^\dagger$. Quantum theory is mainly concerned with normal matrices of which Hermitian ($\mathbf{H}^\dagger = \mathbf{H}$) and unitary ($\mathbf{U}^\dagger = \mathbf{U}^{-1}$) matrices are the most common. However, we need to learn to deal with the "abnormal" matrices, as well.

Matrix products and eigensolutions for polarizer-counter arrangements

Consider a 45° tilted ($\theta_1 = \beta_1/2 = \pi/4$ or $\beta_1 = 90^\circ$) sheet of polarizer lying below a y-sheet, that is, a $\beta_1 = 90^\circ$ (45°-polarized) filter followed by a $\beta_2 = 0^\circ$ filter with bottom path open which is like a y-polarized sheet.



(The $\beta_2 = 0^\circ$ filter with bottom path open is the same as a $\beta_2 = 180^\circ$ filter with top path open.) The transfer matrices for these filters are those of projection operators for the states they let pass. (Recall (2.1.22).)

$$T(2) = |y\rangle\langle y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad T(1) = |x'\rangle\langle x'| = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (\text{for } |x'\rangle: \theta_2 = \beta_2/2 = 45^\circ)$$

Their matrix product is the transfer matrix for the total system of two filters.

$$T(\text{total}) = T(2) \cdot T(1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

The product's eigenvectors determine the eigenstates or "own-states" that may pass with only a change of overall magnitude or phase. The secular equation yields two eigenvalues and two projectors.

$$\lambda^2 - \frac{1}{2}\lambda + 0 = 0, \text{ or: } \lambda = 0, \frac{1}{2}, \text{ gives projectors } P_0 = \begin{pmatrix} \frac{-1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} - \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \quad P_{1/2} = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} - 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

The first eigenket is the non-zero column of P_0 , namely, $|0\rangle = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ with zero eigenvalue, that is, zero transfer.

The second eigenket is the non-zero column $|\frac{1}{2}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ of $P_{1/2}$ with eigenvalue $1/2$, giving 50% amplitude. The

figure above shows the latter eigenstate, namely y-polarization ($\beta_{IN} = 180^\circ$), entering on the right then emerging on the left with its amplitude cut in half and the probability cut by $(1/2)^2 = 1/4$ or 25%, according to

$$T(\text{total}) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \text{ The eigenbras, namely } \langle 0| = \begin{pmatrix} 1 & 0 \end{pmatrix} \text{ and } \langle \frac{1}{2}| = \begin{pmatrix} 1 & 1 \end{pmatrix} \text{ describe what gets through going the}$$

opposite way, that is, left-to-right, according to equations $\langle 0|T(\text{total}) = 0$ and $\langle \frac{1}{2}|T(\text{total}) = \frac{1}{2} \langle \frac{1}{2}|$. These equations

say that x-polarization gets stopped from going in the left end while 45° polarization would have a 50% transfer. The matrices $T(\beta = 180^\circ)$ and $T(\beta = 90^\circ)$ do not commute. Reversing their order gives a different product and a different set of eigensolutions. In this case, since the factor matrices are Hermitian, reversal would simply interchange the eigenbras with the eigenkets, that is, a †-operation.

(2) Projector completeness and spectral decomposition

The normalized projection operators derive from matrix operator \mathbf{M} and its eigenvalues ϵ_k .

$$\mathbf{P}_k = \frac{\prod_{j \neq k} (\mathbf{M} - \epsilon_j \mathbf{1})}{\prod_{j \neq k} (\epsilon_k - \epsilon_j)} \tag{3.1.15a}$$

The projectors \mathbf{P}_k have been shown to be *eigenoperators* for \mathbf{M} .

$$\mathbf{P}_k \mathbf{M} = \mathbf{M} \mathbf{P}_k = \epsilon_k \mathbf{P}_k \tag{3.1.15b}$$

They have also been shown to satisfy *projector orthonormality and idempotency*.

$$\mathbf{P}_j \mathbf{P}_k = \delta_{jk} \mathbf{P}_k \tag{3.1.15c}$$

Now we will demonstrate that they also satisfy a *projector completeness relation*

$$\mathbf{1} = \mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_n \tag{3.1.15d}$$

and a very powerful relation called *spectral decomposition of an operator M*.

$$\mathbf{M} = \epsilon_1 \mathbf{P}_1 + \epsilon_2 \mathbf{P}_2 + \dots + \epsilon_n \mathbf{P}_n \tag{3.1.15e}$$

The completeness relation (3.1.15d) resembles the abstraction (2.1.20) of *axiom-4* repeated here.

$$\mathbf{1} = \sum_{k=1}^n \mathbf{P}_k = \sum_{k=1}^n |k\rangle\langle k| \tag{3.1.16}$$

The similarity is no accident, but there is a logical difference between (3.1.15d) and (3.1.16). The latter is a *physical axiom* of quantum wave path completeness, while the former is a *algebraic theorem* being proved here. The spectral decomposition relation (3.1.15e) follows by operating on the completeness relation (3.1.15d) with the matrix \mathbf{M} using its eigen-operator relation (3.1.15b) that is an algebraic result.

First, let us check that (3.1.15d) and (3.1.16) are correct for examples (3.1.14) expressed in terms of an *outer* or *Kronecker tensor* (\otimes) *product* of eigen-bras and kets from "bad cop" matrix \mathbf{K} .

$$\mathbf{P}_1 = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} = k_1 \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{2} \end{pmatrix} \otimes \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} / k_1 = |\epsilon_1\rangle\langle\epsilon_1|, \quad \mathbf{P}_2 = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} = k_2 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \otimes \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \end{pmatrix} / k_2 = |\epsilon_2\rangle\langle\epsilon_2|, \tag{3.1.14}_{repeat}$$

This agrees with the original results in (3.1.11)_{example}. Furthermore, they sum up to $\mathbf{1}$ as required. More importantly, they provide the following spectral decomposition (3.1.15e) of \mathbf{K} .

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \tag{3.1.15d}_{example} \quad \mathbf{K} = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} = 1 \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + 5 \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \tag{3.1.15e}_{example}$$

Spectral decomposition quickly finds the 100th power \mathbf{K}^{100} of \mathbf{K} using (3.1.15e).

$$\mathbf{K}^{100} = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}^{100} = 1^{100} \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix} + 5^{100} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1+3 \cdot 5^{100} & 5^{100}-1 \\ -3+3 \cdot 5^{100} & 5^{100}+3 \end{pmatrix}$$

Idempotence and orthonormality (3.1.15c) kills all cross terms so a function $f(\mathbf{M})$ of a matrix \mathbf{M} reduces to a sum of projectors weighted with the function evaluated at \mathbf{M} -eigenvalues ϵ_k .

$$f(\mathbf{M}) = f(\epsilon_1) \mathbf{P}_1 + f(\epsilon_2) \mathbf{P}_2 + \dots + f(\epsilon_n) \mathbf{P}_n \tag{3.1.17}$$

This is a *functional spectral decomposition of an operator M*. (Try \mathbf{K}^{-1} , or $\sqrt{\mathbf{K}}$ to test this technique.)

Now to prove the matrix completeness relation (3.1.15d) we will appeal to the numerical analysis lore. The formula (3.1.17) for functional spectral decomposition resembles the terms in the famous *Lagrange interpolation formula* of a function $f(x)$ approximated by its value at N discrete points.

$$L(f(x)) = \sum_{m=1}^N f(x_m) \cdot P_m(x) \quad \text{where: } P_m(x) = \frac{\prod_{j \neq m}^N (x - x_j)}{\prod_{j \neq m}^N (x_m - x_j)} \tag{3.1.18}$$

Lagrange’s formula fits a polynomial of degree $N - 1$ to N arbitrary points $\{x_1, x_2, \dots, x_n\}$ on a function curve $y = f(x)$. Note that each polynomial term $P_m(x)$ has zeros at each point $x = x_j$ *except* $x = x_m$ where $P_m(x_m) = 1$. So at each of these points x_m this L-approximation becomes exact: $L(f(x_m)) = f(x_m)$.

If $f(x)$ happens to be a polynomial of degree $N - 1$ or less, then the L-approximation is exact everywhere, that is, $L(f(x)) = f(x)$ for *all* points x . This is true since one point determines a constant, two points uniquely determine a line, three points uniquely determine a parabola, and N points uniquely determine an $(N - 1)^{th}$ degree curve. Hence if $(N > 1)$ the following special cases of a constant ($f(x) = 1$) and a line ($f(x) = x$) are exactly determined by $L(f(x)) = f(x)$ for *all* points x .

$$1 = \sum_{m=1}^N P_m(x), \qquad x = \sum_{m=1}^N x_m P_m(x)$$

The first corresponds to matrix completeness (3.1.15d) and the second one to spectral decomposition (3.1.15e).

Now a matrix \mathbf{M} and its powers \mathbf{M}^n obey the same algebra as a simple variable x and its powers x^n . So completeness relation is proved. Furthermore, it is true *for all distinct values of the eigenvalue parameters* $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$. Completeness relation (3.1.15d) seems to be *more than true!* This is easily seen for $N = 2$.

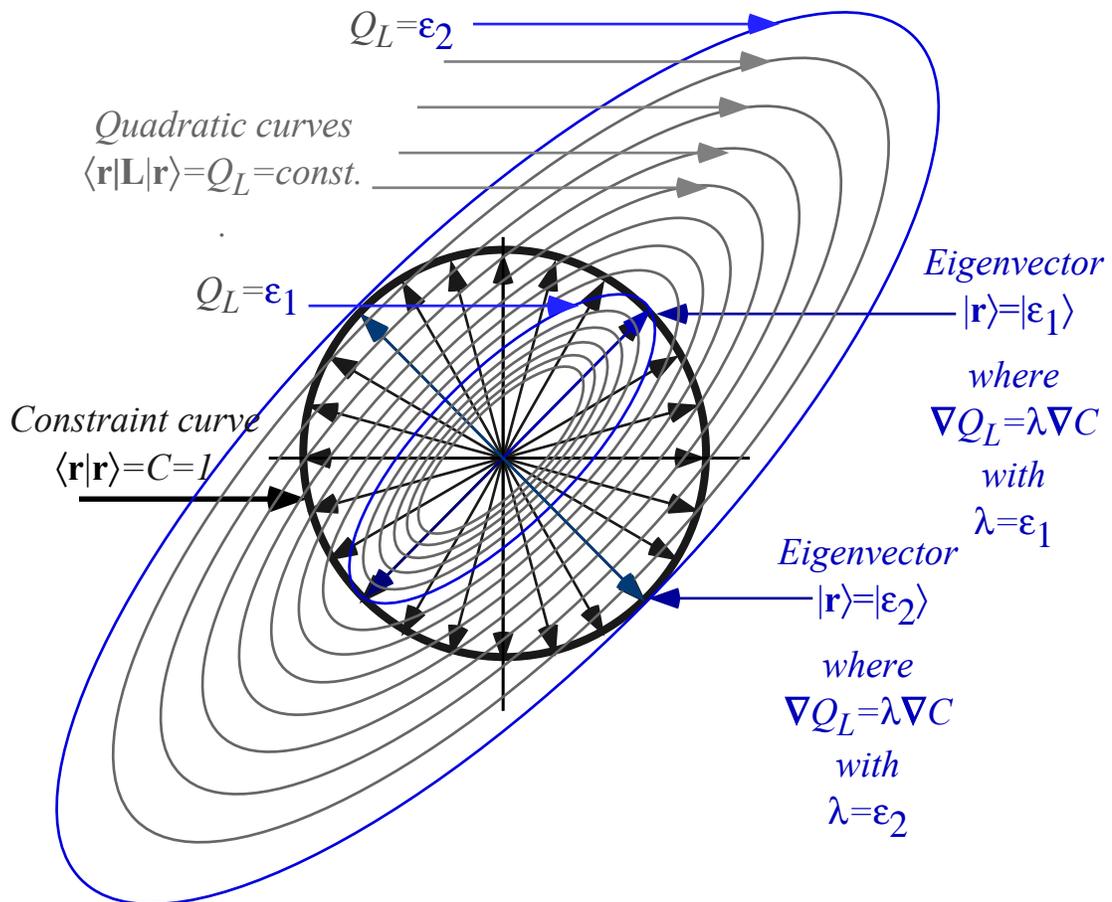
$$\mathbf{P}_1 + \mathbf{P}_2 = \frac{\prod_{j \neq 1} (\mathbf{M} - \varepsilon_j \mathbf{1})}{\prod_{j \neq 1} (\varepsilon_1 - \varepsilon_j)} + \frac{\prod_{j \neq 2} (\mathbf{M} - \varepsilon_j \mathbf{1})}{\prod_{j \neq 2} (\varepsilon_2 - \varepsilon_j)} = \frac{(\mathbf{M} - \varepsilon_2 \mathbf{1})}{(\varepsilon_1 - \varepsilon_2)} + \frac{(\mathbf{M} - \varepsilon_1 \mathbf{1})}{(\varepsilon_2 - \varepsilon_1)} = \frac{(\mathbf{M} - \varepsilon_2 \mathbf{1}) - (\mathbf{M} - \varepsilon_1 \mathbf{1})}{(\varepsilon_1 - \varepsilon_2)} = \frac{-\varepsilon_2 \mathbf{1} + \varepsilon_1 \mathbf{1}}{(\varepsilon_1 - \varepsilon_2)} = \mathbf{1} \quad (\text{for all } \varepsilon_j)$$

Direct algebraic verification of universal identity of completeness (3.1.15d) grows algebraically laborious for larger $N = 3, 4, \dots$, so it is nice that the Lagrangian analogy shows its ε_k independence for all N .

However, the ε_k are required to be the correct eigenvalues ε_k of matrix \mathbf{M} for the other relations in (3.1.15) such as orthogonality and spectral theorems.

Eigensolutions are stationary or extreme-value solutions

Eigenvalues λ of a matrix \mathbf{L} can be viewed as stationary-values of its quadratic form $Q_{\mathbf{L}}(\mathbf{r}) = \langle \mathbf{r} | \mathbf{L} | \mathbf{r} \rangle$, that is, the min-max values of the function $Q(\mathbf{r})$ subject to the constraint of unit norm: $C(\mathbf{r}) = \langle \mathbf{r} | \mathbf{r} \rangle = 1$. Multi-dimensional constrained min-max problems may be solved using *Lagrange multiplier theory* as we will sketch here. The idea is to find those values of $Q_{\mathbf{L}}$ and vector \mathbf{r} for which the $Q_{\mathbf{L}}(\mathbf{r}) = \text{const.}$ ellipses is tangent to the constraint circle $C(\mathbf{r}) = 1$. Stated another way, imagine walking around the constraint circle $C(\mathbf{r}) = \langle \mathbf{r} | \mathbf{r} \rangle = 1$ in the figure below and looking for those places where one of the $Q_{\mathbf{L}}(\mathbf{r}) = \text{const.}$ ellipses is tangent to the unit circle $C(\mathbf{r}) = 1$. Lagrange pointed out that such points would have the gradient $\nabla Q_{\mathbf{L}}$ pointing in the same direction as ∇C , that is, the two gradient vectors $\nabla Q_{\mathbf{L}}$ and ∇C would be proportional to each. In honor of Lagrange, the proportionality constant is taken to be λ in $\nabla Q_{\mathbf{L}} = \lambda \nabla C$, where λ is called a *Lagrange Multiplier*.



The fact that λ symbolizes both the eigenvalue and a Lagrange multiplier is no coincidence; they are equal here. The gradients $\nabla Q_{\mathbf{L}} = \langle \mathbf{r} | \mathbf{L} + \mathbf{L} | \mathbf{r} \rangle$ and $\nabla C = \langle \mathbf{r} | + | \mathbf{r} \rangle$ in Lagrange equation give eigenvalue equations.

$$\mathbf{L} | \mathbf{r} \rangle = \lambda | \mathbf{r} \rangle \text{ and } \langle \mathbf{r} | \mathbf{L} = \langle \mathbf{r} | \lambda$$

On the eigen-directions the Lagrange multiplier is also the value of the quadratic form: $\lambda = Q_{\mathbf{L}}(\mathbf{r}) = \langle \mathbf{r} | \mathbf{L} | \mathbf{r} \rangle$

$$\text{for: } | \mathbf{r} \rangle = | \epsilon_1 \rangle, Q_{\mathbf{L}}(\mathbf{r}) = \langle \epsilon_1 | \mathbf{L} | \epsilon_1 \rangle = \epsilon_1, \quad \text{and for: } | \mathbf{r} \rangle = | \epsilon_2 \rangle, Q_{\mathbf{L}}(\mathbf{r}) = \langle \epsilon_2 | \mathbf{L} | \epsilon_2 \rangle = \epsilon_2.$$

$\langle \mathbf{r} | \mathbf{L} | \mathbf{r} \rangle$ is called an *expectation value* of matrix \mathbf{L} at \mathbf{r} . Eigenvalues are extreme expectation values.

(3) Diagonalizing transformations from projectors

The real goal of many quantum problems is to find a *d-tran matrix*, the transformation matrix that diagonalizes some quantum analyzer matrix **T** or other types of matrices not yet discussed like a scattering matrix **S** or a Hamiltonian matrix **H**. If all (meaning a complete set) of the projection matrices **P_k** are known for a given matrix **M** then the *diagonalization transformation (d-tran)* matrix is easy to get.

Here's how. First you use the columns of the **P_k** matrices to give a set of normalized eigenket vectors. This was described in the preceding section using an asymmetric ("bad cop") matrix **K** as an example and will be discussed further in the following sections. Then you load these columns into the d-tran matrix in whatever order you find convenient. That's all there is to it.

Consider that old "bad cop" matrix **K** again. Since it is asymmetric it requires an extra step you won't need for quantum matrices, but it's instructive to see this, too, given our eigenvectors (3.1.14) in Fig. 3.1.3.

$$\left\{ \langle \varepsilon_1 | = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \langle \varepsilon_2 | = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \end{pmatrix} \right\}, \left\{ | \varepsilon_1 \rangle = \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{2} \end{pmatrix}, | \varepsilon_2 \rangle = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \right\}$$

$(\varepsilon_1, \varepsilon_2) \leftarrow (1,2)$ *d-Tran matrix* $(1,2) \leftarrow (\varepsilon_1, \varepsilon_2)$ *INVERSE d-Tran matrix*

$$\begin{pmatrix} \langle \varepsilon_1 | 1 \rangle & \langle \varepsilon_1 | 2 \rangle \\ \langle \varepsilon_2 | 1 \rangle & \langle \varepsilon_2 | 2 \rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \langle 1 | \varepsilon_1 \rangle & \langle 1 | \varepsilon_2 \rangle \\ \langle 2 | \varepsilon_1 \rangle & \langle 2 | \varepsilon_2 \rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$$

Load distinct bras $\langle \varepsilon_1 |$ and $\langle \varepsilon_2 |$ into d-tran **rows**, and load kets $| \varepsilon_1 \rangle$ and $| \varepsilon_2 \rangle$ into inverse d-tran **columns**. It helps to use *Dirac labeling for all components* so actual transformation is done correctly as shown below.

$$\begin{pmatrix} \langle \varepsilon_1 | 1 \rangle & \langle \varepsilon_1 | 2 \rangle \\ \langle \varepsilon_2 | 1 \rangle & \langle \varepsilon_2 | 2 \rangle \end{pmatrix} \cdot \begin{pmatrix} \langle 1 | \mathbf{K} | 1 \rangle & \langle 1 | \mathbf{K} | 2 \rangle \\ \langle 2 | \mathbf{K} | 1 \rangle & \langle 2 | \mathbf{K} | 2 \rangle \end{pmatrix} \cdot \begin{pmatrix} \langle 1 | \varepsilon_1 \rangle & \langle 1 | \varepsilon_2 \rangle \\ \langle 2 | \varepsilon_1 \rangle & \langle 2 | \varepsilon_2 \rangle \end{pmatrix} = \begin{pmatrix} \langle \varepsilon_1 | \mathbf{K} | \varepsilon_1 \rangle & \langle \varepsilon_1 | \mathbf{K} | \varepsilon_2 \rangle \\ \langle \varepsilon_2 | \mathbf{K} | \varepsilon_1 \rangle & \langle \varepsilon_2 | \mathbf{K} | \varepsilon_2 \rangle \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

It is a good idea to check that your inverse-d-tran is really the inverse of your d-tran.

$$\begin{pmatrix} \langle \varepsilon_1 | 1 \rangle & \langle \varepsilon_1 | 2 \rangle \\ \langle \varepsilon_2 | 1 \rangle & \langle \varepsilon_2 | 2 \rangle \end{pmatrix} \cdot \begin{pmatrix} \langle 1 | \varepsilon_1 \rangle & \langle 1 | \varepsilon_2 \rangle \\ \langle 2 | \varepsilon_1 \rangle & \langle 2 | \varepsilon_2 \rangle \end{pmatrix} = \begin{pmatrix} \langle \varepsilon_1 | \mathbf{1} | \varepsilon_1 \rangle & \langle \varepsilon_1 | \mathbf{1} | \varepsilon_2 \rangle \\ \langle \varepsilon_2 | \mathbf{1} | \varepsilon_1 \rangle & \langle \varepsilon_2 | \mathbf{1} | \varepsilon_2 \rangle \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

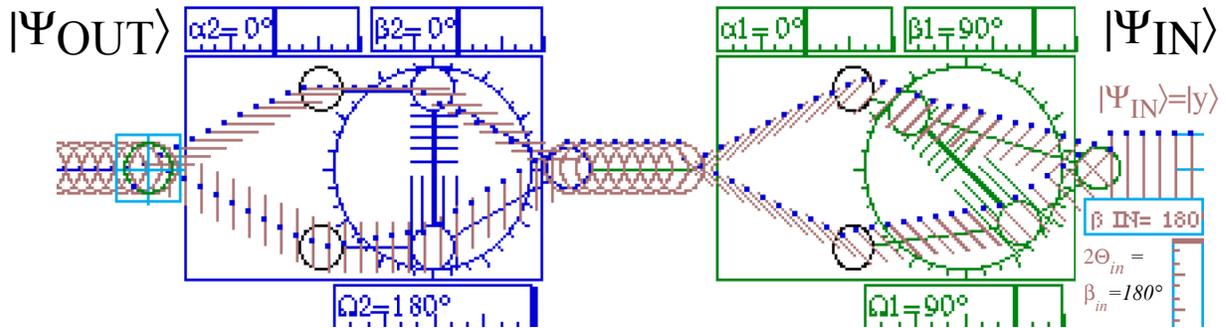
In standard quantum matrices, inverse of a d-tran matrix is its Hermitian conjugate (\dagger). (Not so above!)

$$\begin{pmatrix} \langle \varepsilon_1 | 1 \rangle & \langle \varepsilon_1 | 2 \rangle \\ \langle \varepsilon_2 | 1 \rangle & \langle \varepsilon_2 | 2 \rangle \end{pmatrix} = \begin{pmatrix} \langle 1 | \varepsilon_1 \rangle & \langle 1 | \varepsilon_2 \rangle \\ \langle 2 | \varepsilon_1 \rangle & \langle 2 | \varepsilon_2 \rangle \end{pmatrix}^\dagger = \begin{pmatrix} \langle 1 | \varepsilon_1 \rangle^* & \langle 2 | \varepsilon_1 \rangle^* \\ \langle 1 | \varepsilon_2 \rangle^* & \langle 2 | \varepsilon_2 \rangle^* \end{pmatrix} = \begin{pmatrix} \langle 1 | \varepsilon_1 \rangle & \langle 1 | \varepsilon_2 \rangle \\ \langle 2 | \varepsilon_1 \rangle & \langle 2 | \varepsilon_2 \rangle \end{pmatrix}^{-1}$$

The outer matrices are equal in all cases, even for "bad cop" matrix **K**. The inner matrix definition of (\dagger) can be set equal to the outer ones for unitary or Hermitian matrices like the "good cop" example **H**.

Matrix products and eigensolutions for active analyzers

Consider a 45° tilted ($\theta_1 = \beta_1/2 = \pi/4$ or $\beta_1 = 90^\circ$) analyzer followed by a untilted ($\beta_2 = 0$) analyzer shown below. Active analyzers have both paths open and a phase shift $e^{-i\Omega}$ between each path as in the examples introduced in Fig. 1.3.5. Here the first analyzer has $\Omega_1 = 90^\circ$. The second has $\Omega_2 = 180^\circ$.



The transfer matrix for each analyzer is a sum of projection operators for each open path multiplied by the phase factor that is active at that path. Here we will simply apply the entire phase factor $e^{-i\Omega_1} = e^{-i\pi/2}$ to the top path in the first analyzer and the factor $e^{-i\Omega_2} = e^{-i\pi}$ to the top path in the second analyzer.

$$T(2) = e^{-i\pi} |x\rangle\langle x| + |y\rangle\langle y| = \begin{pmatrix} e^{-i\pi} & 0 \\ 0 & 1 \end{pmatrix} \quad T(1) = e^{-i\pi/2} |x'\rangle\langle x'| + |y'\rangle\langle y'| = e^{-i\pi/2} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1-i}{2} & \frac{-1-i}{2} \\ -\frac{1-i}{2} & \frac{1-i}{2} \end{pmatrix}$$

The matrix product $T(total) = T(2)T(1)$ relates input states $|\Psi_{IN}\rangle$ to output states: $|\Psi_{OUT}\rangle = T(total)|\Psi_{IN}\rangle$

$$T(total) = T(2)T(1) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1-i}{2} & \frac{-1-i}{2} \\ -\frac{1-i}{2} & \frac{1-i}{2} \end{pmatrix} = \begin{pmatrix} \frac{-1+i}{2} & \frac{1+i}{2} \\ -\frac{1-i}{2} & \frac{1-i}{2} \end{pmatrix} = e^{-i\pi/4} \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \sim \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

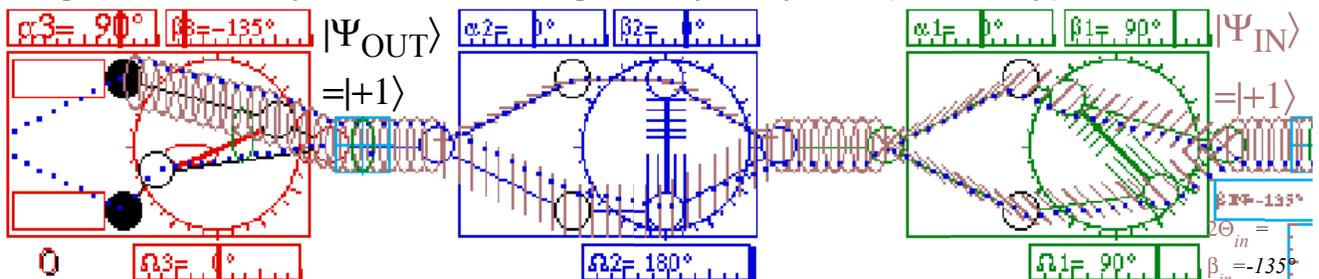
We drop the overall phase $e^{-i\pi/4}$ since it is unobservable. $T(total)$ yields two eigenvalues and projectors.

$$\lambda^2 - 0\lambda - 1 = 0, \text{ or: } \lambda = +1, -1$$

, gives projectors

$$P_{+1} = \frac{\begin{pmatrix} \frac{-1}{\sqrt{2}} + 1 & \frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} + 1 \end{pmatrix}}{1 - (-1)} = \frac{\begin{pmatrix} -1 + \sqrt{2} & i \\ -i & 1 + \sqrt{2} \end{pmatrix}}{2\sqrt{2}}, \quad P_{-1} = \frac{\begin{pmatrix} 1 + \sqrt{2} & -i \\ i & -1 + \sqrt{2} \end{pmatrix}}{2\sqrt{2}}$$

The first eigenvector $|+1\rangle = \begin{pmatrix} -1 + \sqrt{2} \\ -i \end{pmatrix} = \begin{pmatrix} 0.414 \\ -i \end{pmatrix}$ is a vertical left-handed ellipse with ratio $x:y = 0.414:1$. This eigen-ellipse must exit analyzer-2 as the same ellipse. Analyzer-3 yields $\alpha_3 = 90^\circ$ and $\beta_3 = -135^\circ$.



The other eigenvector $|-1\rangle$ is horizontal right-handed ellipse with inverse ratio $x:y = 1:2.414$ and angles $\alpha_3 = 90^\circ$ and $\beta_3 = 45^\circ$. The meaning of the electron spin angles α and β is described in section 2.10.

(c) Eigenvector projectors (Degenerate eigenvalues)

We have just shown that any matrix with distinct eigenvalues can be spectrally decomposed, i.e., diagonalized. What if the secular equation (3.1.5d) of an N -by- N matrix \mathbf{H} has some degenerate eigenvalues $\varepsilon_1 = \varepsilon_2 \dots = \varepsilon_{\ell_1}$? If so, it is possible that \mathbf{H} cannot be completely diagonalized, though this is rarely the case. It all depends upon whether or not the HC equation (3.1.8) really needs its repeated factors. Suppose each eigenvalue ε_j is ℓ_j -fold degenerate so the secular equation factors as follows:

$$S(\varepsilon) = 0 = (-1)^N (\varepsilon - \varepsilon_1)^{\ell_1} (\varepsilon - \varepsilon_2)^{\ell_2} \dots (\varepsilon - \varepsilon_p)^{\ell_p} \quad (3.1.19a)$$

where $\ell_1 + \ell_2 + \dots + \ell_p = N$. Then the N -th degree HC equation is:

$$\mathbf{0} = (-1)^N (\mathbf{H} - \varepsilon_1 \mathbf{1})^{\ell_1} (\mathbf{H} - \varepsilon_2 \mathbf{1})^{\ell_2} \dots (\mathbf{H} - \varepsilon_p \mathbf{1})^{\ell_p} \quad (3.1.19b)$$

Each eigenvalue ε_j is repeated ℓ_j times as is each factor $(\mathbf{H} - \varepsilon_j \mathbf{1})$ in the HC equation. The number ℓ_j is called the *degree of degeneracy* of eigenvalue ε_j .

Suppose, now you find that only *one* of each distinct factor is needed to give a *matrix* zero, that is, the following p -th degree equation holds.

$$\mathbf{0} = (\mathbf{H} - \varepsilon_1 \mathbf{1})(\mathbf{H} - \varepsilon_2 \mathbf{1}) \dots (\mathbf{H} - \varepsilon_p \mathbf{1}) \quad (3.1.20)$$

This is just like the distinct eigenvalue situation in equation (3.1.8), so the matrix \mathbf{H} is completely diagonalizable and spectrally decomposable using the same techniques described previously.

(1) Minimal equation and diagonalizability criterion

Otherwise, if \mathbf{H} does not satisfy a non-degenerate equation then it is not diagonalizable. The lowest degree polynomial equation a matrix \mathbf{H} can satisfy is called its *minimal* equation. (If all roots are distinct, that is $p=N$, then the HC-equation is the minimal equation.)

When only one of each of p distinct factors $(\mathbf{H} - \varepsilon_j \mathbf{1})$ in the minimal polynomial is needed to give zero, then removing that factor gives p non-zero $(p-1)$ -th degree operators $\{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_p\}$ following (3.1.15a). They are idempotent ($\mathbf{P}_j^2 = \mathbf{P}_j$), orthogonal ($\mathbf{P}_i \mathbf{P}_j = \mathbf{0}$ if $i \neq j$) and complete ($\sum_{j=1}^p \mathbf{P}_j = \mathbf{1}$) just as in the case of no repeated roots. Here is the key *diagonalizability criterion*.

In general, an orthogonal and complete set of \mathbf{P}_j 's is possible, if and only if, the \mathbf{H} minimal equation has no repeated factors. Then and only then is matrix \mathbf{H} diagonalizable.

(2) Nilpotent operators ("Bad" degeneracy)

Repeated $(\mathbf{H} - \varepsilon_j \mathbf{1})$ factors in the minimal equation are always fatal for the process of building a complete set of idempotents \mathbf{P}_j . Even one repeat is fatal, suppose:

$$\mathbf{0} = (\mathbf{H} - \varepsilon_1 \mathbf{1})^2 (\mathbf{H} - \varepsilon_2 \mathbf{1}) \dots, \text{ but: } \mathbf{N} = (\mathbf{H} - \varepsilon_1 \mathbf{1})(\mathbf{H} - \varepsilon_2 \mathbf{1}) \dots \neq \mathbf{0} \quad (3.1.21)$$

Removal of one repeat gives a non-zero operator \mathbf{N} whose square has the missing $(\mathbf{H} - \varepsilon_1 \mathbf{1})$ that gives zero.

$$\mathbf{N}^2 = (\mathbf{H} - \varepsilon_1 \mathbf{1})^2 (\mathbf{H} - \varepsilon_2 \mathbf{1})^2 \dots = \mathbf{0}$$

(The presence of additional commuting factors $(\mathbf{H} - \varepsilon_2 \mathbf{1}) \dots$ does not save it.) Such an operator is called a *nilpotent operator* or, simply a nilpotent. A nilpotent is a troublesome and 'unwanted beast' for the basic diagonalization process but an essential feature of Non-Abelian symmetry analysis.

For example, consider a 'bad' degenerate matrix. (This is not just a "bad cop" but a real "crook"!)

$$\mathbf{B} = \begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix}$$

Its secular equation has two equal roots ($\varepsilon = b$ twice).

$$S(\varepsilon) = \varepsilon^2 - 2b\varepsilon + b^2 = (\varepsilon - b)^2 = 0$$

The HC equation is then as follows.

$$S(\mathbf{B}) = \mathbf{B}^2 - 2b\mathbf{B} + b^2\mathbf{1} = (\mathbf{B} - b\mathbf{1})^2 = \mathbf{0}$$

The matrix factor

$$\mathbf{N} = \mathbf{B} - b\mathbf{1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (3.1.22)$$

is an example of nilpotent eigen-projector which satisfies

$$\mathbf{N}^2 = \mathbf{0} \text{ (but } \mathbf{N} \neq \mathbf{0} \text{) and: } \mathbf{BN} = b\mathbf{N} = \mathbf{NB}. \quad (3.1.23)$$

The nilpotent contains only one non-zero eigenket and one eigenbra.

$$|b\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \langle b| = (0 \ 1) \quad (3.1.24)$$

Also, they are orthogonal to each other! ($\langle b|b\rangle = 0$) There can be no completeness, orthonormality, spectral decomposition or diagonalization for this 'bad' degenerate matrix in the ordinary sense of (3.1.17).

Let us not give the impression that nilpotents or other "bad" matrices are not valuable for general quantum theory. In fact the operator described in (3.1.22) is an example of an *elementary operator* \mathbf{e}_{ab}

$$\mathbf{e}_{12} = |1\rangle\langle 2|$$

Along with its partners it makes up a 4-dimensional (recall Sec. 2.2d) *U(2) unit tensor operator space*

$$U(2) \text{ op-space} = \{\mathbf{e}_{11} = |1\rangle\langle 1|, \mathbf{e}_{12} = |1\rangle\langle 2|, \mathbf{e}_{21} = |2\rangle\langle 1|, \mathbf{e}_{22} = |2\rangle\langle 2|\} \quad (3.1.25a)$$

out of which all *U(2)* operators are made by linear combination. They obey a *simple matrix algebra*

$$\mathbf{e}_{ij} \mathbf{e}_{km} = \delta_{jk} \mathbf{e}_{im} \quad (3.1.25b)$$

This is very useful stuff later on. Just be aware you cannot diagonalize an \mathbf{e}_{ab} for $a \neq b$!

(3) Multiple diagonalization ("Good" degeneracy)

An example of a 'good' degenerate (but still diagonalizable) matrix is the anti-diagonal "gamma" matrix \mathbf{G} which Dirac used to generate Lorentz transformations.

$$\mathbf{G} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (3.1.26)$$

It has a 4th degree secular equation.

$$S(\varepsilon) = 0 = \varepsilon^4 - 2\varepsilon^2 + 1 = (\varepsilon - 1)^2 (\varepsilon + 1)^2$$

There are two pairs of degenerate roots ($\varepsilon = \pm 1$, twice), but \mathbf{G} satisfies only a second degree minimal equation. (Check this!)

$$\mathbf{0} = (\mathbf{G} - \mathbf{1})(\mathbf{G} + \mathbf{1}) \quad (3.1.27)$$

This allows us to use theory based on projection formula (3.1.15) to derive two projection operators.

$$\mathbf{P}_1^G = \frac{\mathbf{G} - (-1)\mathbf{1}}{1 - (-1)} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad (3.1.28a)$$

$$\mathbf{P}_{-1}^G = \frac{\mathbf{G} - (1)\mathbf{1}}{-1 - (1)} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \quad (3.1.28b)$$

These satisfy all orthonormality and completeness or spectral decomposition relations (3.1.15a-d).

The main difference here is that each of these projectors contains two linearly independent ket vectors: from the first and second columns of \mathbf{P}_1 we get $|1_1\rangle$ and $|1_2\rangle$, and from \mathbf{P}_{-1} we get $|-1_1\rangle$ and $|-1_2\rangle$. (Recall that we showed in (3.1.20)_{example H} that each \mathbf{P}_j contains all the scalar products and normalization constants of its bra-rows and ket-columns.)

$$|1_1\rangle = \frac{|1_1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad |1_2\rangle = \frac{|1_2\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad |-1_1\rangle = \frac{|-1_1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \quad |-1_2\rangle = \frac{|-1_2\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \quad (3.1.29)$$

This example is particularly convenient since the $(1|2)$ components $(\mathbf{P}_j)_{12}$ happen to be zero, and therefore first and second rows are already orthogonal ($(j_1|j_2) = 0$). Otherwise we would need to orthogonalize to get a second orthonormal eigenket. Such a process is called *Gram-Schmidt orthogonalization* which is described below.

Gram-Schmidt orthogonalization

Suppose a non-zero scalar product $(\langle j_1 | j_2 \rangle \neq 0)$ exists between two vectors. Then we would need to orthogonalize to get a second orthonormal eigenket

$$|j_{\text{normal } 2}\rangle = N_1 |j_1\rangle + N_2 |j_2\rangle \quad (3.1.30a)$$

such that

$$\begin{aligned} \langle j_1 | j_{\text{normal } 2} \rangle &= 0 = N_1 \langle 1|1 \rangle + N_2 \langle 1|2 \rangle \\ \langle j_{\text{normal } 2} | j_{\text{normal } 2} \rangle &= 1 = N_1^2 \langle 1|1 \rangle + N_1 N_2 (\langle 1|2 \rangle + \langle 2|1 \rangle) + N_2^2 \langle 2|2 \rangle \end{aligned}$$

As we noted the a-row and b-column scalar product matrix is just the \mathbf{P}_j matrix, itself.

$$(a|b) = (\mathbf{P}_j)_{ab}$$

$(a|b)$ is sometimes called a *Grammian* matrix. Solving for (3.1.30a) coefficients gives

$$N_1 = -N_2 \frac{\langle 1|2 \rangle}{\langle 1|1 \rangle} \quad \text{where } N_2 = \sqrt{\frac{1}{\langle 2|2 \rangle - \frac{\langle 1|2 \rangle \langle 2|1 \rangle}{\langle 1|1 \rangle}}} \quad (3.1.30b)$$

This Gram Schmidt orthonormalization (3.1.30) is not a unique solution since any linear combination of degenerate eigenvectors is still an eigenvector. To help sort this out we consider below a more elegant procedure using spectral decomposition.

(d) Projector splitting: A key to algebraic reduction

Dirac notation for the \mathbf{G} example completeness relation using eigenvectors (3.1.29) is the following:

$$1 = \mathbf{P}_1^G + \mathbf{P}_{-1}^G = |1_1\rangle\langle 1_1| + |1_2\rangle\langle 1_2| + |-1_1\rangle\langle -1_1| + |-1_2\rangle\langle -1_2| \quad (3.1.31a)$$

$$= \mathbf{P}_{1_1} + \mathbf{P}_{1_2} + \mathbf{P}_{-1_1} + \mathbf{P}_{-1_2} \quad (3.1.32b)$$

Here the original projection operators (3.1.28) have each been “split” in two.

$$\begin{aligned} \mathbf{P}_1^G = \mathbf{P}_{1_1} + \mathbf{P}_{1_2} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &= |1_1\rangle\langle 1_1| + |1_2\rangle\langle 1_2| \end{aligned} \quad (3.1.32c)$$

$$\begin{aligned} \mathbf{P}_{-1}^G = \mathbf{P}_{-1_1} + \mathbf{P}_{-1_2} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &= |-1_1\rangle\langle -1_1| + |-1_2\rangle\langle -1_2| \end{aligned} \quad (3.1.32d)$$

Idempotent *projector splitting*, such as (3.1.32c-d), is an important process in the application of symmetry groups to quantum theory. Our first examples are the completeness splitting of the unit operator $\mathbf{1}$. Let us now see the power of splitting algebra and an important technique in symmetry analysis.

Suppose we are given two mutually commuting matrix operators: the \mathbf{G} from (3.1.26) before, and another operator \mathbf{H} .

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \quad (3.1.33)$$

(First, it is important to verify that they do, in fact, commute.)

$$\mathbf{GH} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \mathbf{HG} \quad (3.1.34)$$

This implies that the projection \mathbf{P}_g^G operators (3.1.28) of \mathbf{G} commute with any new projection operators \mathbf{P}_k^H generated by \mathbf{H} . This will lead to a combined set \mathbf{P}^{GH} which simultaneously spectrally decomposes both \mathbf{G} and \mathbf{H} . The new \mathbf{P}_k^H operators follow from the secular and minimal equations for \mathbf{H} .

$$\begin{aligned} \mathbf{P}_2^H &= \frac{(\mathbf{H} - (-2)\mathbf{1})}{2 - (-2)} & \mathbf{P}_{-2}^H &= \frac{(\mathbf{H} - (2)\mathbf{1})}{-2 - (2)} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} & &= \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \end{aligned} \quad (3.1.35a)$$

These obey the following completeness relations and spectral decomposition for \mathbf{G} and \mathbf{H} , separately.

$$\mathbf{1} = \mathbf{P}_1^G + \mathbf{P}_{-1}^G \quad \mathbf{1} = \mathbf{P}_2^H + \mathbf{P}_{-2}^H \quad (3.1.35b)$$

$$\mathbf{G} = (1)\mathbf{P}_1^G + (-1)\mathbf{P}_{-1}^G \quad \mathbf{H} = (2)\mathbf{P}_2^H + (-2)\mathbf{P}_{-2}^H \quad (3.1.35c)$$

The old "1=1·1 trick"

By multiplying the two completeness relations one obtains a set of projectors that, together, satisfy orthonormality (because $\mathbf{P}_j^G \mathbf{P}_k^H = \mathbf{P}_k^H \mathbf{P}_j^G$) and completeness (because $\mathbf{1} = \mathbf{1} \cdot \mathbf{1}$).

$$\mathbf{1} = \mathbf{1} \cdot \mathbf{1} = (\mathbf{P}_1^G + \mathbf{P}_{-1}^G)(\mathbf{P}_2^H + \mathbf{P}_{-2}^H) \quad (3.1.36a)$$

$$\mathbf{1} = (\mathbf{P}_1^G \mathbf{P}_2^H + \mathbf{P}_1^G \mathbf{P}_{-2}^H + \mathbf{P}_{-1}^G \mathbf{P}_2^H + \mathbf{P}_{-1}^G \mathbf{P}_{-2}^H) \quad (3.1.36b)$$

(We call this the *'the old one-equals-one-times-one'* trick!) Matrix multiplication gives four new operators which in this case are orthonormal and complete projectors.

$$\begin{aligned}
 \mathbf{P}_{1,2}^{GH} \equiv \mathbf{P}_1^G \mathbf{P}_2^H = & \quad \mathbf{P}_{1,-2}^{GH} \equiv \mathbf{P}_1^G \mathbf{P}_{-2}^H = & \quad \mathbf{P}_{-1,2}^{GH} \equiv \mathbf{P}_{-1}^G \mathbf{P}_2^H = & \quad \mathbf{P}_{-1,-2}^{GH} \equiv \mathbf{P}_{-1}^G \mathbf{P}_{-2}^H = \\
 \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} & \quad \frac{1}{4} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} & \quad \frac{1}{4} \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix} & \quad \frac{1}{4} \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \\
 & & & (3.1.37a)
 \end{aligned}$$

(This isn't the way it always works out. Some may come out zero, but the ones that are not zero must be orthonormal idempotent projectors.) Each is automatically an eigen-operator of both **G** and **H**. (Note: **H** commutivity is needed!)

$$\mathbf{G}\mathbf{P}_{g,h}^{GH} = \mathbf{G}\mathbf{P}_g^G \mathbf{P}_h^H = \epsilon_g^G \mathbf{P}_{g,h}^{GH} \tag{3.1.37b}$$

$$\mathbf{H}\mathbf{P}_{g,h}^{GH} = \mathbf{H}\mathbf{P}_g^G \mathbf{P}_h^H = \mathbf{P}_g^G \mathbf{H}\mathbf{P}_h^H = \epsilon_h^H \mathbf{P}_{g,h}^{GH} \tag{3.1.37c}$$

Thus, the $\mathbf{P}_{j,k}^{GH}$ spectrally decompose both **G** and **H**, simultaneously.

$$\mathbf{G} = (1)\mathbf{P}_{1,2}^{GH} + (1)\mathbf{P}_{1,-2}^{GH} + (-1)\mathbf{P}_{-1,2}^{GH} + (-1)\mathbf{P}_{-1,-2}^{GH} \tag{3.1.37d}$$

$$\mathbf{H} = (2)\mathbf{P}_{1,2}^{GH} + (-2)\mathbf{P}_{1,-2}^{GH} + (2)\mathbf{P}_{-1,2}^{GH} + (-2)\mathbf{P}_{-1,-2}^{GH} \tag{3.1.37e}$$

So, by simple matrix multiplication we have accomplished an idempotent splitting like that in (3.1.32) without needing to Gram-Schmidt orthogonalize bra-kets. (Yes!)

$$\mathbf{P}_1^G = \mathbf{P}_{1,2}^{GH} + \mathbf{P}_{1,-2}^{GH} \tag{3.1.38a}$$

$$\mathbf{P}_{-1}^G = \mathbf{P}_{-1,2}^{GH} + \mathbf{P}_{-1,-2}^{GH} \tag{3.1.38b}$$

Most important, the splitting is “just right” for the new **H** matrix; finding the “right” Gram-Schmidt combination (3.1.30) to diagonalize both **H** and **G** at once, would require even more calculation.

In this case, no further idempotent splitting of (3.1.36b) is possible. For $N - by - N$ matrix operators there can have no more than N linearly independent eigenvectors and no more than N orthonormal projectors. Each $\mathbf{P}_{g,h}^{GH}$ in (1.2.50c) has in its columns and rows one and only one independent eigenvector. Such an ‘unsplittable’ projector is called an *irreducible* idempotent or projector.

You can tell how many irreducible projectors are "hiding" inside a given idempotent projector **P** (*reducible*) matrix by taking its trace. This *splitting number* is equal to the trace.

$$\text{Splitting number} = \text{Trace}\mathbf{P}(\text{reducible}) = \text{Number of irreducible projectors in } \mathbf{P} \tag{3.1.39}$$

Irreducible projectors have unit trace!

$$\text{Trace}\mathbf{P}(\text{irreducible}) = 1 \tag{3.1.40}$$

Each of (3.1.37c) projectors have a unit trace as they should.

Note that such a complete splitting as we saw in (3.1.37) was not guaranteed. It depends on what **H** operator we chose to do the splitting. It could have happened that one or more of the $\mathbf{P}_g^G \mathbf{P}_h^H = \mathbf{P}_{g,h}^{GH}$ products in (3.1.37b) came out to be zero. Then some of the non-zero $\mathbf{P}_{g,h}^{GH}$ will not be irreducible. Suppose, for example, we chose $\mathbf{H} = 2\mathbf{G}$: then $\mathbf{P}_{1,-2}^{GH}$ and $\mathbf{P}_{-1,2}^{GH}$ are zero while $\mathbf{P}_{1,2}^{GH} = \mathbf{P}_1^G$ and $\mathbf{P}_{-1,-2}^{GH} = \mathbf{P}_{-1}^G$ remain reducible.

Having a set of N irreducible \mathbf{G} and \mathbf{H} projectors like (3.1.37) is useful since any N -by- N operator \mathbf{K} which commutes with \mathbf{G} and \mathbf{H} must share exactly the same N projectors no matter what is the form of \mathbf{K} 's individual completeness relation.

$$\mathbf{1} = \sum_{k=1}^n \mathbf{P}_k^K \quad (n \leq N)$$

To see this note that a product of \mathbf{K} completeness relation by $\mathbf{P}_{g,h}^{GH}$ must have one *and only one* non-zero term.

$$\mathbf{P}_{gh}^{GH} = \sum_{k=1}^N \mathbf{P}_{g,h}^{GH} \mathbf{P}_k^K = \mathbf{P}_{g,h}^{GH} \mathbf{P}_k^K \quad (3.1.41)$$

Since $\mathbf{P}_{g,h}^{GH}$ is irreducible it cannot split into new orthogonal idempotent projectors, and this implies that each $\mathbf{P}_{g,h}^{GH}$ is already an eigen-operator for \mathbf{K} .

$$\mathbf{K} \mathbf{P}_{g,h}^{GH} = \mathbf{P}_{g,h}^{GH} \mathbf{K} \mathbf{P}_{k'}^K = \left(\epsilon_{k'}^K \right) \mathbf{P}_{g,h}^{GH} \quad (3.1.42)$$

We will then have diagonalized \mathbf{K} with less mathematical labor than the old Gram-Schmidt methods.

(e) Why symmetry groups are useful

The results ending with (3.1.42) illustrate an important symmetry technique. Imagine you wanted to diagonalize a complicated matrix \mathbf{K} and knew that it commutes with some other operators \mathbf{G} and \mathbf{H} for which irreducible projectors are more easily found. Then you don't have to bother with the secular equation of \mathbf{K} and may just multiply \mathbf{K} by the projectors provided by \mathbf{G} and \mathbf{H} as in (3.1.42) above.

In later chapters we will see how having a group of operators $\{ \mathbf{G}, \mathbf{H}, \dots \}$ that commute with a big system matrix \mathbf{K} helps to reduce its secular equation and sometimes solve it completely. When transformation operators $\mathbf{G}, \mathbf{H}, \dots$ (like rotations) commute with an analyzer matrix $\mathbf{T}=\mathbf{K}$ (or other type of quantum system matrix \mathbf{K}) it means that

$$\mathbf{K} \mathbf{G} = \mathbf{G} \mathbf{K} \quad \text{or} \quad \mathbf{G}^\dagger \mathbf{K} \mathbf{G} = \mathbf{K} \quad \text{or} \quad \mathbf{G} \mathbf{K} \mathbf{G}^\dagger = \mathbf{K} \quad (3.1.43)$$

which means \mathbf{K} is *invariant* to the transformation induced by \mathbf{G} . This is called a *symmetry* of the system \mathbf{K} stands for and it is often pretty easy to spot. The group of these operators is called a *symmetry group*.

Entire groups can be spectrally decomposed into irreducible projection operators, and then these can be used to decompose the system matrix \mathbf{K} into one set of \mathbf{P} 's made of \mathbf{G}, \mathbf{H} , symmetry operators.

More to the point, because \mathbf{K} is a spectral combination (3.1.15c) of \mathbf{P} 's and \mathbf{P} 's are in turn combinations (3.1.15a) or (3.1.37) of powers and products of $\mathbf{G}, \mathbf{H}, \dots$ it follows that \mathbf{K} is a *linear combination of its own group of symmetry operators*, including $\mathbf{G}, \mathbf{H}, \dots$ and their products. This is a very powerful idea! It will be useful in some problems and then be used extensively following Unit 3.

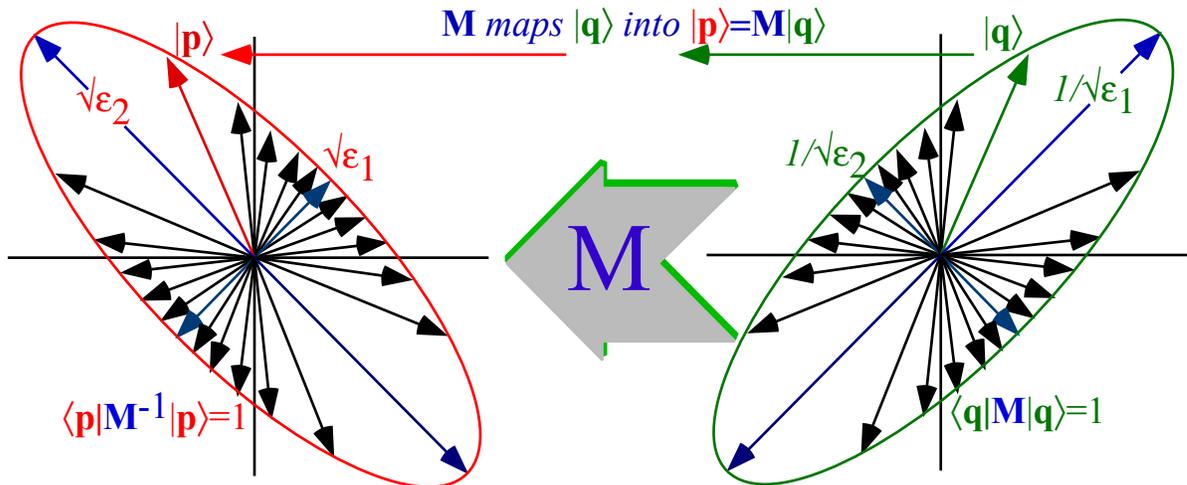
To summarize, we use the spectral decomposition of some easily "killed" operators to attack more difficult ones, much as a "killed" virus in a vaccine saves us from suffering troublesome or dangerous diseases.

Quadratic surfaces help to visualize matrix operations

The mapping $M|c\rangle=|r\rangle$ of a unit circle $\langle c|c\rangle=1$ by symmetric matrix M is an ellipse $I=\langle r|M^{-2}|r\rangle$ as shown in a previous sidebar. If, instead of mapping vectors $|c\rangle$ on a circle, we map vectors $|q\rangle$ on a surface corresponding to a unit constant *quadratic form* $I = \langle q|M|q\rangle$, the resulting vectors $|p\rangle = M|q\rangle$ of this mapping will lie on a related quadratic surface given by

$$I = \langle q|M|q\rangle = \langle q|p\rangle = \langle p|M^{-1}|p\rangle$$

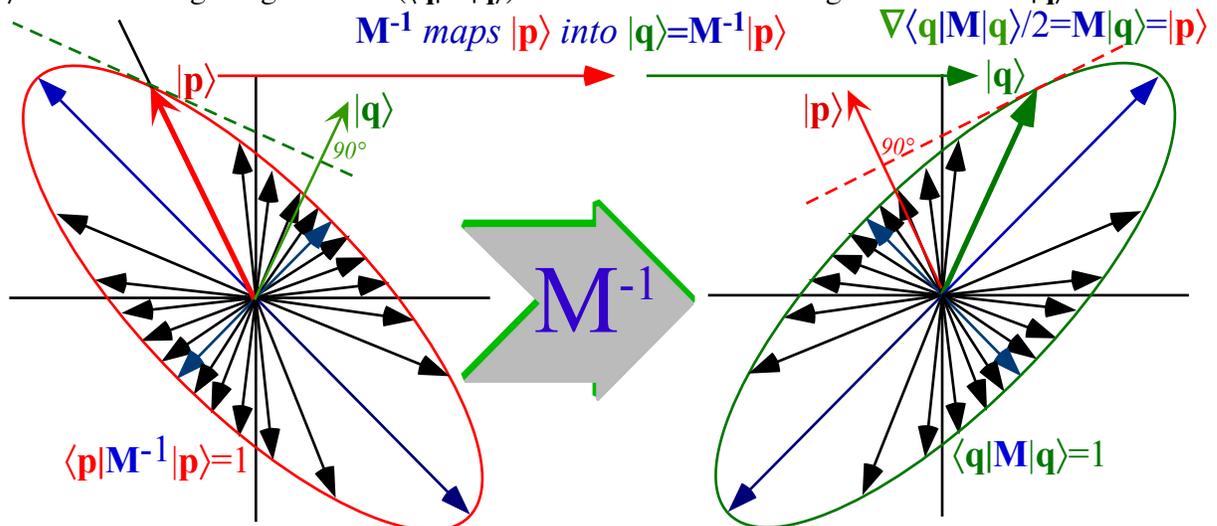
The surface $I = \langle p|M^{-1}|p\rangle$ defined by vectors $|p\rangle$ is called the conjugate or *inverse quadratic form*. An example of such a mapping is displayed in the figure below. The semi-axes of the $|p\rangle$ ellipse are square roots of eigenvalues $\sqrt{\epsilon_1}$ and $\sqrt{\epsilon_2}$ while $|q\rangle$ ellipse axes are inverse roots $1/\sqrt{\epsilon_1}$ and $1/\sqrt{\epsilon_2}$.



The precise geometry of this mapping is found by considering the gradient of the quadratic curves.

$$\nabla(\langle q|M|q\rangle)=\langle q|M + M|q\rangle = 2 M|q\rangle = 2 |p\rangle$$

Let matrix M be real symmetric so there is no distinction between bras and kets. This shows that the mapped vector $|p\rangle$ must lie along the gradient $\nabla(\langle q|M|q\rangle)$ that is normal to the tangent to curve at $|q\rangle$.



The inverse map works in the same way since $|q\rangle$ is normal to the tangent at mapped point $|p\rangle$. It should be noted that quadratic surfaces can be hyperbolic as well as elliptic if there are negative eigenvalues. Eigenvectors are any vectors that are in the same direction as quadratic curve gradient at their point.

3.2 Approximate Eigensolutions by Perturbation Techniques

One of the alternatives to numerical diagonalization or symmetry analysis techniques is called *perturbation analysis*. This often is a viable alternative for problems with little or no symmetry because such problems usually do not have resonances or degeneracies that often come with having symmetry. Then eigenvalues and vectors may change by only tiny amounts that can be approximated.

Perturbation techniques, like most "approximologies" are many and varied. Their use can be more art than a science. We discuss one here based upon analysis of the secular determinant (3.1.5d).

$$0 = \det|\mathbf{H} - \lambda\mathbf{1}| = \det \begin{vmatrix} H_{11} - \lambda & H_{12} & H_{13} & H_{14} & \cdots \\ H_{21} & H_{22} - \lambda & H_{23} & H_{24} & \cdots \\ H_{31} & H_{32} & H_{33} - \lambda & H_{34} & \cdots \\ H_{41} & H_{42} & H_{43} & H_{44} - \lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix} = \det \begin{vmatrix} D_{11} & D_{12} & D_{13} & D_{14} & \cdots \\ D_{21} & D_{22} & D_{23} & D_{24} & \cdots \\ D_{31} & D_{32} & D_{33} & D_{34} & \cdots \\ D_{41} & D_{42} & D_{43} & D_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

where:

$$D_{\mu\nu} = \begin{cases} H_{\mu\mu} - \lambda & \text{if: } \mu=\nu \\ H_{\mu\nu} & \text{if: } \mu \neq \nu \end{cases} = H_{\mu\nu} - \delta_{\mu\nu}\lambda \quad (3.2.1)$$

(a) Secular determinantal expansion

The ϵ -tensor sum of permutations reviewed in Appendix 3.A-B. We do the sum class by class since each class of permutation is either all even (+) or all odd (-). We'll show only the terms for $N=4$.

$$0 = \det|\mathbf{H} - \lambda\mathbf{1}| = \sum_{perm} \epsilon_{\alpha\beta\gamma\delta} \dots D_{1\alpha} D_{2\beta} D_{3\gamma} D_{4\delta} \dots \quad (3.2.2)$$

First there is the "zero-flip" term corresponding to partition 1+1+1+1...

$$0 = \det|\mathbf{H} - \lambda\mathbf{1}| = D_{11} D_{22} D_{33} D_{44} \dots = (H_{11} - \lambda)(H_{22} - \lambda)(H_{33} - \lambda)(H_{44} - \lambda) \dots \quad (3.2.3a)$$

Then we subtract (odd) "one-flip" terms corresponding to partition 2+1+1... (There are $N(N-1)/2$ of these)

$$\begin{aligned} \dots - I(12) - I(13) - I(14) \dots &= \dots - H_{12} H_{21} D_{33} D_{44} - H_{13} D_{22} H_{31} D_{44} - H_{14} D_{22} D_{33} H_{41} \dots \\ - I(23) - I(24) \dots &\quad - D_{11} H_{23} H_{32} D_{44} - D_{11} H_{24} D_{33} H_{42} \dots \\ - I(34) \dots &\quad - D_{11} D_{22} H_{34} H_{43} \dots \end{aligned} \quad (3b)$$

Add the "two-flip" terms corresponding to partition 3+1... (There are $N(N-1)(N-2)/3$ of these.) Recall that (143) = (314) = (431) means, "1 goes where 4 was, 4 goes where 3 was, and 3 goes where 1 was," and the inverse is (143)⁻¹ = (134) = (341) = (413). It's called "two-flip" because (abc) = (ac)(bc) is two flips.

$$\begin{aligned} \dots + I(123) + I(124) + I(134) + I(234) \dots &= \dots + H_{13} H_{21} H_{32} D_{44} + H_{14} H_{21} D_{33} H_{42} + H_{14} D_{22} H_{31} H_{43} + D_{11} H_{24} H_{32} H_{43} \dots \\ \dots + I(132) + I(142) + I(143) + I(243) \dots &= \dots + H_{12} H_{23} H_{31} D_{44} + H_{12} H_{24} D_{33} H_{41} + H_{13} D_{22} H_{34} H_{41} + D_{11} H_{23} H_{34} H_{42} \dots \end{aligned} \quad (3c)$$

Add the "two-flip" terms corresponding to partition 2+2... (There are $N(N-1)(N-2)(N-3)/8$ of these.)

$$\dots + I(12)(34) + I(13)(24) + I(14)(23) \dots = \dots + H_{12} H_{21} H_{34} H_{43} + H_{13} H_{24} H_{31} H_{42} + H_{14} H_{23} H_{32} H_{41} + \dots \quad (3d)$$

Finally, (if N were really 4) subtract (odd) "three-flip" terms. (There are $N(N-1)(N-2)(N-3)/4$ of these.)

$$\begin{aligned} & \cdot - I(1234) - I(1324) - I(1423) \dots = \cdot - H_{14} H_{22} H_{32} H_{43} - H_{14} H_{23} H_{31} H_{42} - H_{13} H_{24} H_{32} H_{41} \dots \\ & \cdot - I(1432) - I(1423) - I(1324) \dots = \cdot - H_{12} H_{23} H_{34} H_{41} - H_{13} H_{24} H_{32} H_{41} - H_{14} H_{23} H_{31} H_{42} \dots \end{aligned} \quad (3e)$$

Collect these results and replace diagonal D_{mm} factors with $H_{mm} - \lambda$.

$$0 = (H_{11} - \lambda)(H_{22} - \lambda)(H_{33} - \lambda)(H_{44} - \lambda) \dots \quad (3.2.4a)$$

and include the $N(N-1)/2=6$ terms for "one-flip" partition 2+1+1...

$$\begin{aligned} \dots - H_{12} H_{21} (H_{33} - \lambda)(H_{44} - \lambda) - H_{13} (H_{22} - \lambda) H_{31} (H_{44} - \lambda) - H_{14} (H_{22} - \lambda)(H_{33} - \lambda) H_{41} \dots \\ - (H_{11} - \lambda) H_{23} H_{32} (H_{44} - \lambda) - (H_{11} - \lambda) H_{24} (H_{33} - \lambda) H_{42} \dots \\ - (H_{11} - \lambda)(H_{22} - \lambda) H_{34} H_{43} \dots \end{aligned} \quad (4b)$$

and the $N(N-1)(N-2)/3=8$ terms for "two-flip" partition 3+1...

$$\begin{aligned} \dots + H_{13} H_{21} H_{32} (H_{44} - \lambda) + H_{14} H_{21} (H_{33} - \lambda) H_{42} + H_{14} (H_{22} - \lambda) H_{31} H_{43} + (H_{11} - \lambda) H_{24} H_{32} H_{43} \dots \\ + H_{12} H_{23} H_{31} (H_{44} - \lambda) + H_{12} H_{24} (H_{33} - \lambda) H_{41} + H_{13} (H_{22} - \lambda) H_{34} H_{41} + (H_{11} - \lambda) H_{23} H_{34} H_{42} \dots \end{aligned} \quad (4c)$$

and the $N(N-1)(N-2)(N-4)/8=3$ terms for the other "two-flip" partition 2+2...

$$\dots + H_{12} H_{21} H_{34} H_{43} + H_{13} H_{24} H_{31} H_{42} + H_{14} H_{23} H_{32} H_{41} + \dots \quad (4.d)$$

and, finally the $N(N-1)(N-2)(N-4)/4=6$ terms (for $N=4$) "three-flip" partition 4...

$$\begin{aligned} \dots - H_{14} H_{22} H_{32} H_{43} - H_{14} H_{23} H_{31} H_{42} - H_{13} H_{24} H_{32} H_{41} \dots \\ - H_{12} H_{23} H_{34} H_{41} - H_{13} H_{24} H_{32} H_{41} - H_{14} H_{23} H_{31} H_{42} \dots \end{aligned} \quad (4.e)$$

(b) Perturbation approximations

Now we look at the art of approximation. Suppose we want to approximate the one unknown eigenvalue $\lambda = E_I$ closest to the known \mathbf{H} -matrix diagonal element H_{11} . Suppose further that all the other diagonal differences $|H_{11} - H_{22}|$, $|H_{22} - H_{33}|$, $|H_{33} - H_{44}|$, ... are larger than the magnitudes off-diagonal matrix elements H_{12} , H_{13} , ..., H_{24} , H_{34} , etc. Then we can divide the secular equation by the large factors $(H_{22} - \lambda)$ $(H_{33} - \lambda)$ $(H_{44} - \lambda)$ and leave behind the (supposedly) small factor $(H_{11} - \lambda)$. Then we collect terms on the right hand side $(H_{11} - \lambda)$ terms that can be discarded since they should be tiny.

$$\begin{aligned} 0 = (H_{11} - \lambda) & - \frac{H_{12} H_{21}}{(H_{22} - \lambda)} - \frac{H_{13} H_{31}}{(H_{33} - \lambda)} - \frac{H_{14} H_{41}}{(H_{44} - \lambda)} \dots & - \frac{(H_{11} - \lambda) H_{23} H_{32}}{(H_{22} - \lambda)(H_{33} - \lambda)} - \frac{(H_{11} - \lambda) H_{24} H_{42}}{(H_{22} - \lambda)(H_{44} - \lambda)} \\ & \left(\Updownarrow \text{"keepers"} \Updownarrow \right) & \left(\Updownarrow \text{"discards"} \Updownarrow \Rightarrow \right) - \frac{(H_{11} - \lambda) H_{34} H_{43}}{(H_{33} - \lambda)(H_{44} - \lambda)} \\ \dots + \frac{H_{12} H_{23} H_{31} + H_{13} H_{32} H_{21}}{(H_{22} - \lambda)(H_{33} - \lambda)} + \frac{H_{12} H_{24} H_{41} + H_{14} H_{42} H_{21}}{(H_{22} - \lambda)(H_{44} - \lambda)} \dots & + \frac{(H_{11} - \lambda)(H_{23} H_{34} H_{42} + H_{24} H_{43} H_{32})}{(H_{22} - \lambda)(H_{33} - \lambda)(H_{44} - \lambda)} \\ & + \frac{H_{13} H_{34} H_{41} + H_{14} H_{43} H_{31}}{(H_{33} - \lambda)(H_{44} - \lambda)} + \dots \\ \dots + \frac{H_{12} H_{21} H_{34} H_{43} + H_{13} H_{24} H_{31} H_{42} + H_{14} H_{23} H_{32} H_{41}}{(H_{22} - \lambda)(H_{33} - \lambda)(H_{44} - \lambda)} + \dots \end{aligned}$$

$$\frac{-H_{14}H_{43}H_{32}H_{22} - H_{14}H_{42}H_{23}H_{31} - H_{13}H_{32}H_{24}H_{41}\dots}{(H_{22} - \lambda)(H_{33} - \lambda)(H_{44} - \lambda)}$$

$$\frac{-H_{12}H_{23}H_{34}H_{41} - H_{13}H_{32}H_{24}H_{41} - H_{14}H_{42}H_{23}H_{31}\dots}{(H_{22} - \lambda)(H_{33} - \lambda)(H_{44} - \lambda)}$$

Now we have an equation for the unknown perturbed eigenvalue λ with one more approximation, that is, to replace every λ in the "keeper" denominators by the approximate energy $\lambda \sim H_{11} = E_1^{(0)}$. (In fact, this substitution kills the "discard" terms.)

$$\lambda = H_{11} - \frac{H_{12}H_{21}}{(H_{22} - H_{11})} - \frac{H_{13}H_{31}}{(H_{33} - H_{11})} - \frac{H_{14}H_{41}}{(H_{44} - H_{11})} \dots$$

$$\dots + \frac{H_{12}H_{23}H_{31} + H_{13}H_{32}H_{21}}{(H_{22} - H_{11})(H_{33} - H_{11})} + \frac{H_{12}H_{24}H_{41} + H_{14}H_{42}H_{21}}{(H_{22} - H_{11})(H_{44} - H_{11})} + \frac{H_{13}H_{34}H_{41} + H_{14}H_{43}H_{31}}{(H_{33} - H_{11})(H_{44} - H_{11})} \dots$$

The terms that are fourth order in H_{mn} are left off above, but included below in the final result.

$$\lambda = E_1 + \sum_{j \neq 1}^N \frac{H_{1j}H_{j1}}{(E_1 - E_j)} + \sum_{j \neq 1}^N \sum_{k \neq 1, j}^N \frac{H_{1j}H_{jk}H_{k1}}{(E_1 - E_j)(E_1 - E_k)} + \sum_{j=1}^N \sum_{k \neq j}^N \sum_{\ell \neq j, k}^N \frac{H_{1j}H_{jk}H_{k\ell}H_{\ell 1}}{(E_1 - E_j)(E_1 - E_k)(E_1 - E_\ell)} \dots$$

(3.2.5a)

Here the diagonal terms are denoted as approximate eigenvalues:

$$E_m = H_{mm} \tag{3.2.5b}$$

A diagrammatic representation of this is given in Fig. 3.2.1. Note that the choice of the number-1 value is arbitrary. This approximation works just as well replacing 1 by 2, 3, ..., or N . The figure indicates number 1 as the lowest eigenvalue but that is not a necessary condition, either. Neither are all the terms in the order chosen for the figure; it shows only one of many combinations and permutations of the 3rd and 4th order terms.

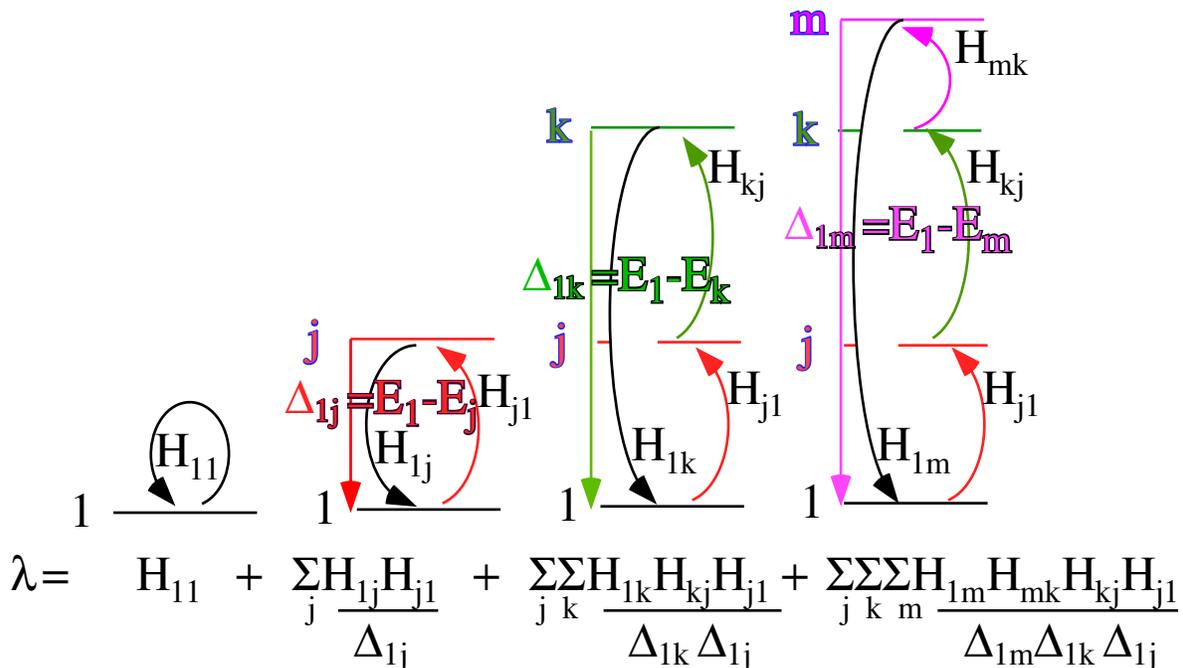


Fig. 3.2.1 Diagrammatic description of perturbation series

Later, in Chapter 9, we will show how the eigenvalues of the Hamiltonian energy matrix \mathbf{H} correspond to quantized energy levels E_j and energy eigenstates $|\epsilon_j\rangle$. According to this we interpret each perturbation term as tracing a path or circuit between the approximate eigenvalues $E_j = H_{jj}$ which correspond to, as yet, imprecisely defined energy states $|\epsilon_j\rangle$ which are initially nothing but the original base states $|\epsilon_j\rangle \sim |j\rangle$ for the problem.

Each path begins and ends on the level that one is interested in defining more precisely. (In Fig. 3.2.1 it is called level E_l .) The path visits a number of intermediate levels $E_m = H_{mm}$ once (and only once) and each one has what is called an *energy* or *resonance denominator*

$$\Delta_m = E_l - E_m = H_{ll} - H_{mm} \quad (3.2.6)$$

This determines, along with matrix element products $H_{km}H_{mj}$, a contributing factor $H_{km}H_{mj} / \Delta_m$ for the intermediate base state $|\epsilon_m\rangle \sim |m\rangle$ to the energy correction for that path. Obviously, a zero or near-zero energy denominator Δ_m would signal a major or infinite contribution of one path and one base state. Unfortunately, it would also signal the invalidity of the perturbation approximation.

(c) Testing perturbation approximation with exact 2x2 eigenvalues

In order to see how well perturbation theory works, it helps to compare its lowest order predictions with that of direct and exact diagonalization. By choosing a simple two-by-two matrix such as

$$\mathbf{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} E_1 & V \\ V & E_2 \end{pmatrix}, \quad (3.2.7)$$

this test is easy to do. We choose off-diagonal V to be real to make it even easier.

First, do the perturbation calculation. (It's often the first thing to try!) Up to second order we have the following approximate eigenvalues using (3.2.5) to the second order term.

$$\begin{aligned} \lambda_1 &= E_1 + \frac{V^2}{E_1 - E_2}, \\ \lambda_2 &= E_2 + \frac{V^2}{E_2 - E_1}. \end{aligned} \quad (3.2.8)$$

Then the exact calculation starts with the secular equation (3.1.5).

$$\lambda^2 - (\text{Trace}\mathbf{H})\lambda + \det|\mathbf{H}| = 0 = \lambda^2 - (E_1 + E_2)\lambda + (E_1E_2 - V^2) \quad (3.2.9a)$$

The two roots are

$$\lambda_{1,2} = \frac{E_1 + E_2 \pm \sqrt{(E_1 + E_2)^2 - 4E_1E_2 + 4V^2}}{2} = \frac{E_1 + E_2 \pm \sqrt{(E_1 - E_2)^2 + 4V^2}}{2}, \quad (3.2.9b)$$

The comparison is made by assuming (as in (3.2.5)) that V is small compared to $|E_1 - E_2|$. Then the binomial approximation $(a+b)^{1/2} \sim a^{1/2} + b/(2a^{1/2})$...gives

$$\lambda_{1,2} = 1/2 \left(E_1 + E_2 \pm (E_1 - E_2) \pm 4V^2 / (2(E_1 - E_2)) \right), \quad (3.2.9c)$$

which agrees perfectly with second order perturbation approximation (3.2.8).

The results are plotted in Fig. 3.2.2 below to show the differences. The 2nd order perturbation approximation fits a parabola to the exact hyperbola of each eigenvalue trace versus the off-diagonal element V . As expected, the perturbation approximation deviates as the off-diagonal matrix element V increases. However, it would improve with an increase in the difference $|E_1 - E_2|$, the two are related and it is a larger ratio of V to $|E_1 - E_2|$ that will make a perturbation approximation less accurate.

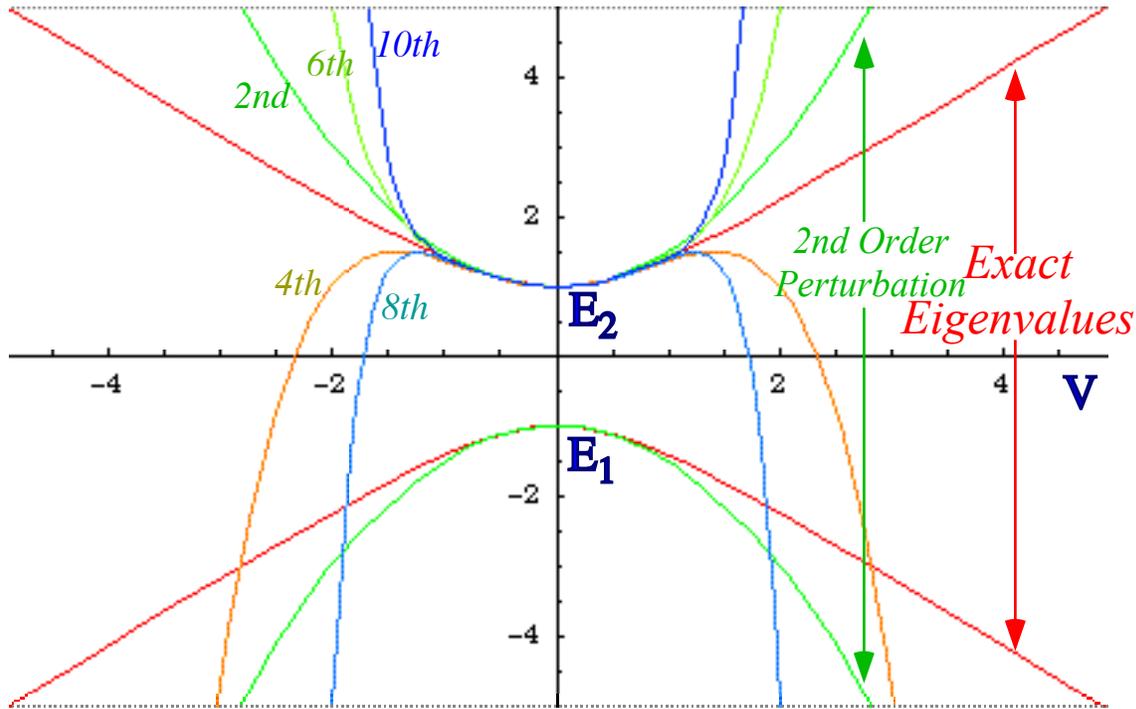


Fig. 3.2.2 Comparison of exact vs. 2nd-order thru 8th-order perturbation approximations

Second order perturbation formulas are simple, easy to apply, and, for this example, at least, quite an effective approximation for a range of V roughly equal to $|E_1 - E_2|$. The same cannot be said for higher order perturbation terms, particularly those of 6th or higher which seem to follow a law of diminishing returns. Even a 10th order formula only extends the range of validity a little in Fig. 3.2.2. Worse, a simple application of (3.2.5) to the two-level problem is wrong for 6th and higher orders. A direct application of (3.2.5) gives

$$E_2 = \frac{\Delta}{2} + \frac{V^2}{\Delta} - \frac{V^4}{\Delta^3} + \frac{V^6}{\Delta^5} - \frac{V^8}{\Delta^7} + \frac{V^{10}}{\Delta^9} \dots, \text{ where: } \Delta = |E_1 - E_2|$$

while the correct binomial expansion of the exact result (3.2.9) which is plotted in Fig. 3.2.2 is

$$E_2 = \frac{\Delta}{2} + \frac{V^2}{\Delta} - \frac{V^4}{\Delta^3} + \frac{2V^6}{\Delta^5} - \frac{5V^8}{\Delta^7} + \frac{14V^{10}}{\Delta^9} \dots$$

But, even the corrected polynomials are miserable approximations to the hyperbola approaching its asymptote. Also, the series is divergent. Similar problems exist for 3, 4, ..., or N-level systems.

The problem is that eigenvalues are generally more like oscillatory (sinusoidal) or exponential and hyperbolic functions and do not take kindly to being represented by polynomials. Check this out by comparing a sine wave to its Taylor series polynomial approximation. How many orders do you need to approximate one full oscillation to 1% or better? And, note what happens outside that range of validity!

Appendix 3.A Matrix Determinants, Adjuncts, and Inverses

Determinants of an N -by- N matrix can be dealt with conveniently using the N -th value Levi-Civita ϵ -symbol defined below:

$$\epsilon_{i_1 i_2 i_3 \dots i_N} = \begin{cases} 0: & \text{if any two } i_a \text{ are equal} \\ 1: & \text{if } \{i_1 \dots i_N\} \text{ is EVEN shuffle of } \{1 \dots N\} \\ -1: & \text{if } \{i_1 \dots i_N\} \text{ is ODD shuffle of } \{1 \dots N\} \end{cases} \quad (3.A.1)$$

Then the determinant may be written as a sum over all N^N combination of N integers $\{ i_1 i_2 \dots i_N \}$ between 1 and N .

$$\det |M| = \sum_{\{ i_1 \dots i_N \}} \epsilon_{i_1 i_2 i_3 \dots i_N} M_{1i_1} M_{2i_2} M_{3i_3} \dots M_{Ni_N} \quad (3.A.2)$$

Only $N!$ of these terms actually exist. The non-zero ones are just permutations of $\{ i_1 = 1 i_2 = 2 i_3 = 3 \dots i_N = N \}$. Negative (positive) terms belong to odd (even) permutations. (See Appendix 3.B.)

From now on let us imply a sum (1 -to- N) over any indices-repeated on *only* one side of an equation so we will drop the Σ sign. This is called the *dummy index* sum convention.

The concept of *minor* or *adjunct* component expansions follows easily. Pulling the first component out of (3.A.2) gives (with our sum convention)

$$\det |M| = M_{1i_1} M_{i_1 1}^{ADJ} = M_{11} M_{11}^{ADJ} + M_{12} M_{21}^{ADJ} \dots + M_{1N} M_{N1}^{ADJ}$$

where adjunct components M^{ADJ} are defined below.

$$\begin{aligned} M_{a1}^{ADJ} &= \epsilon_{a i_2 i_3 \dots} M_{2i_2} M_{3i_3} \dots \\ &= -\epsilon_{i_2 a i_3 \dots} M_{2i_2} M_{3i_3} \dots \\ &= \epsilon_{i_2 i_3 a \dots} M_{2i_2} M_{3i_3} \dots \end{aligned} \quad (3.A.2)$$

The adjunct component M_{ab}^{ADJ} is just $(-1)^{a+b}$ times the determinant made after crossing out the b -th row and a -th column of matrix M , and it goes to the a -th row and b -th column of the adjunct matrix M^{ADJ} . The determinant $\det|M|$ equals the matrix product of any row of M and the *same* column of M^{ADJ} .

$$\det|M| = M_{1i_1} M_{i_1 1}^{ADJ} = M_{2i_2} M_{i_2 2}^{ADJ} = \dots \quad (3.A.3)$$

where:

$$M_{i_1 1}^{ADJ} = \varepsilon_{i_1 i_2 i_3} M_{2i_2} M_{3i_3} \dots$$

$$M_{i_2 2}^{ADJ} = \varepsilon_{i_1 i_2 i_3} M_{1i_1} M_{3i_3} \dots$$

⋮

A matrix inverse formula follows by showing that the following matrix product involving, for example, the first row of M and the second column of M^{ADJ} is zero.

$$\begin{aligned} M_{1i_1} M_{i_1 2}^{ADJ} &= \varepsilon_{i_1 i_2 i_3} \dots M_{1i_1} M_{1i_2} M_{3i_3} \dots \\ &= -\varepsilon_{i_2 i_1 i_3} \dots M_{1i_1} M_{1i_2} M_{3i_3} \dots && \text{(Switch two } \varepsilon \text{ indices)} \\ &= -\varepsilon_{i_1 i_2 i_3} \dots M_{1i_2} M_{1i_1} M_{3i_3} \dots && \text{(Relabel two sum indices)} \\ &= -\varepsilon_{i_1 i_2 i_3} \dots M_{1i_1} M_{1i_2} M_{3i_3} \dots = -M_{1i_1} M_{i_1 2}^{ADJ} \\ &= 0 \end{aligned} \quad (3.A.4)$$

Any two equal row factors (the first and second are equal to M_{1i_1} in (3.A.4) above) in the ε -combination makes it vanish due to ε -antisymmetry. So the following general result holds.

$$M_{ai} M_{ib}^{ADJ} = \delta_{ab} \det|M| \quad (3.A.5)$$

So, for non-singular M (non-zero $\det|M|$) the inverse M^{-1} exists and is defined as follows:

$$M_{ab}^{-1} = \frac{M_{ab}^{ADJ}}{\det|M|}, \quad (3.A.6a)$$

so that

$$M_{ai} M_{ib}^{-1} = \delta_{ab}, \quad (3.A.6b)$$

that is, a matrix product of it with M yields a unit matrix ($\mathbf{1} = MM^{-1}$).

A more complete definition of the determinant used ε -tensors on both sides of the equation to reflect the fact that determinants are antisymmetric to column permutations as well as row permutations.

$$\det|M|\varepsilon_{abc\dots} = \varepsilon_{i_1 i_2 i_3 \dots} M_{i_1 a} M_{i_2 b} M_{i_3 c} \dots \quad (3.A.7)$$

This helps to expand matrix products and to prove a useful result: the determinant of a matrix product is simply the product of the determinants of the matrix factors. (Remember: repeated indices are being summed.)

$$\begin{aligned} \det|M \cdot N|\varepsilon_{abc\dots} &= \varepsilon_{i_1 i_2 i_3 \dots} \left(M_{i_1 j_1} N_{j_1 a} \right) \left(M_{i_2 j_2} N_{j_2 b} \right) \left(M_{i_3 j_3} N_{j_3 c} \right) \dots \\ &= \left(\varepsilon_{i_1 i_2 i_3 \dots} M_{i_1 j_1} M_{i_2 j_2} M_{i_3 j_3} \dots \right) N_{j_1 a} N_{j_2 b} N_{j_3 c} \dots \\ &= \det|M|\varepsilon_{j_1 j_2 j_3 \dots} N_{j_1 a} N_{j_2 b} N_{j_3 c} \dots \\ &= \det|M|\det|N|\varepsilon_{abc\dots} \end{aligned} \quad (3.A.8)$$

One corollary of (3.A.6b) and (3.A.8) is the following. (We note $\det|\mathbf{1}|=1$, too.)

$$\det|M| = \frac{1}{\det|M^{-1}|} \quad (3.A.9)$$

Appendix 3.B Classification of Permutations

Suppose there is a neatly ordered set of N billiard balls lined up on a rack according to their numbers $\{1,2,3,4,5,6,7,8,\dots,N\}$. After a game the customers put them back in some permuted order like $\{4,2,8,6,3,7,1,5,\dots,N\}$. (We'll make it simple here and suppose only the first eight balls are out of order.)

Suppose it's your job to straighten them out. You have only two hands so it's natural to switch two at a time. You could look for the 1-ball and switch it with whatever ball is in the number-1 position. In this case the 4-ball is where the 1-ball should be, so you would switch the 1-and-4 balls. Let's write this as an equation using Dirac notation: (Bold numbers indicate which are being switched.)

$$(14)|4,2,8,6,3,7,1,5\rangle = |1,2,8,6,3,7,4,5\rangle \quad (3.B.1)$$

The "2-flip" operation (14) is called a *transposition* or a *2-shuffle* or a *bicycle*. Using only bicycles we can complete the reordering. Looking for the 2-ball we see it's already in the 2nd- position so we don't need to do anything to it. A 'do-nothing' permutation is written as follows.

$$(2)|1,2,8,6,3,7,4,5\rangle = |1,2,8,6,3,7,4,5\rangle$$

The operation (2) is called an *identity transposition* or a *unicycle*. Combining the preceding two equations gives.

$$\begin{aligned} (2)(14)|4,2,8,6,3,7,1,5\rangle &= (2)|1,2,8,6,3,7,4,5\rangle \\ (2)(14)|4,2,8,6,3,7,1,5\rangle &= |1,2,8,6,3,7,4,5\rangle \end{aligned} \quad (3.B.2)$$

Now, the 3-ball needs to go where the 8-ball is currently sitting. So we apply the bicycle (38) to this.

$$(38)(2)(14)|4,2,8,6,3,7,1,5\rangle = (38)|1,2,8,6,3,7,4,5\rangle = |1,2,3,6,8,7,4,5\rangle \quad (3.B.3)$$

Then the 4-ball is put in the 4-th spot where 6-ball was sitting using bicycle (46).

$$(46)(38)(2)(14)|4,2,8,6,3,7,1,5\rangle = (46)|1,2,3,6,8,7,4,5\rangle = |1,2,3,4,8,7,6,5\rangle \quad (3.B.4)$$

Then the 5-ball is put in the 5-th spot where 8-ball was sitting using bicycle (58).

$$(58)(46)(38)(2)(14)|4,2,8,6,3,7,1,5\rangle = (58)|1,2,3,4,8,7,6,5\rangle = |1,2,3,4,5,7,6,8\rangle \quad (3.B.5)$$

Finally, a (67) bicycle finishes the job.

$$(67)(58)(46)(38)(2)(14)|4,2,8,6,3,7,1,5\rangle = |1,2,3,4,5,6,7,8\rangle \quad (3.B.6)$$

Since the whole job took exactly five bicycles this is an *ODD* permutation, and it would get a (-1) sign in an 8-by-8 matrix determinant according to equation (3.A.1). A permutation's *parity* is *EVEN* or *ODD* if it has an even or odd number of bicycles. There are more efficient ways to decompose a permutation but its parity is the same no matter how you do the job.

For example, you may have noticed that we had to move some of the balls more than once. Is there a way to reshuffle while moving each ball just once? The answer is yes if you're able to pick up more than two at a time. This involves permutation *tricycles* (where you pick up three balls at once) or *quadracycles* (where you have to pick up four balls), and so on.

With a little manual and mathematical dexterity we can rewrite the final equation (3.B.6) in a simpler and ultimately more revealing form. First we note that permutation operations commute with each other if they

share no numbers in common. So we can move (46) to the left of (58) and (14) to the left of (2) , (38) , and (58) as follows.

$$(67)(58)(46)(38)(2)(14) = (67)(46)(14) (58)(38) (2) \quad (3.B.7)$$

But, that's as far as you can go since (14) doesn't commute with (46) since both involve the 4-ball. (Try it!)

However, we can combine bicycles that share balls into bigger cycles. For example, two bicycles that share one ball like $(58)(38)$ can be read as follows:

First, ball-3 replaces ball-8. (Right operator (38) acts first.)

Second, ball-8, in turn displaces ball-5. (Left operator (58) acts next.)

Third, ball-5 winds up where ball-8 was after (38) . That's where ball-3 was before (38) .

We write this product as a tricycle

$$(58)(38) = (385) = (538) = (853) \quad (3.B.8)$$

(385) is read as follows: ..3-displaces-8-displaces-5-displaces-3-.... and is the same as ..

(538) which is read: ..5-displaces-3-displaces-8-displaces-5-.... or....

(853) which is read: ..8-displaces-5-displaces-3-displaces-8-.... .

Note that if a bicycle product shares two balls it becomes a unicycle, that is no operation at all!

$$(85)(58) = (58)(58) = (5) = (8) = \dots = (1) \quad (3.B.9)$$

Similarly, a quadracycle is a product of three bicycles such as the following.

$$(67)(46)(14) = (1467) = (4671) = (6714) = (7146) \quad (3.B.10)$$

So our example permutation has 1 bicycle, 1 tricycle, and 1 quadracycle. Not counting the no-op-unicycle, we see that it is done in only two operations instead of five.

$$(67)(58)(46)(38)(2)(14) = (2) (385) (1467) \quad (3.B.11)$$

A graphical example of just such a permutation unraveling is done using a more direct way in Fig. 3.B.1 below. The problem is that it gives the inverse permutation $(1764) (358) (2)$ instead of what we just worked out!

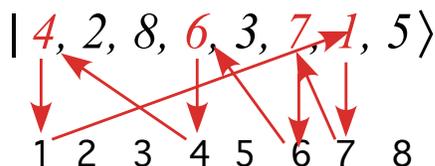
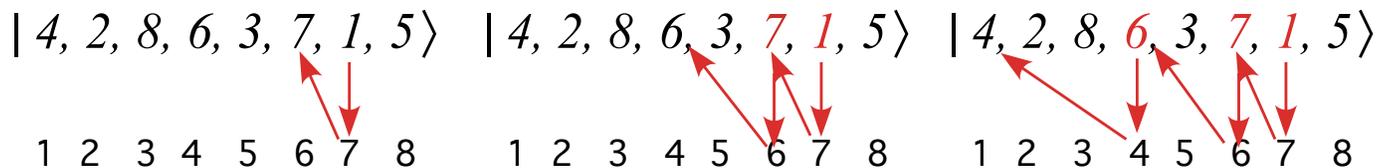
Why?

Welcome to the world of transformation groups! As you will learn if you study this book, every transformation of “things” has to be defined relative to their “pockets.” You may label a transformation using numbers on the things (here, the pool balls) or using numbers on the pockets. As we will see one definition gives the inverse of what the other one gives. This is a very important observation in quantum theory where the “balls” are “particles” and the “pockets” are “states” as will be discussed later.

In the meantime we have already seen a version of this transformational duality in the T-operators or rotation operators that can be defined in “alias” or “alibi” flavors in Section 2.2. A rotation matrix $\langle i|R|k\rangle$ is meaningless unless you specify its bra-kets, that is, its basis. A bra-ket is a two-sided thing, a destination and a point of origin, and all of quantum theory and relativity is concerned with their relative values. Absolutes, one

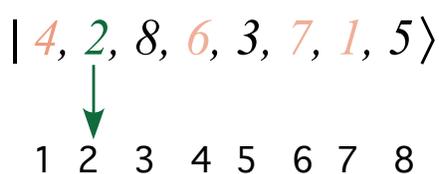
might hope, went out with the absolute monarchs deposed during the 18th century enlightenment.

Unraveling a permutation (Starting with “1”)



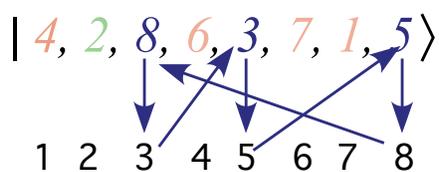
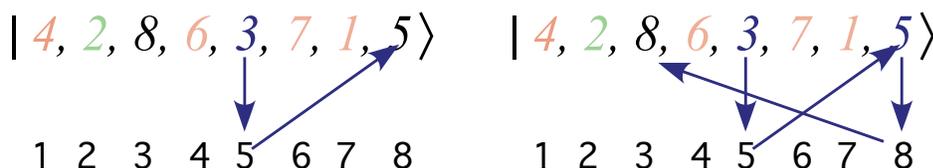
*Closes on a permutation quadracycle
(1764)=(4176)=etc.*

(Next higher number that has not been used is a “2”)



*Closes on a permutation unicycle
(2)*

(Next higher number that has not been used is a “3”)



*Closes on a permutation tricycle
(358)=(835)=etc.*

Final result: (1764)(2)(358)=(358)(1764)(2)=etc.

Fig. 3.B.1 Permutation cycle structure unraveling using pocket numbers.

Permutations are classified by the numbers of v_1 of unicycles, v_2 of bicycles, v_3 of tricycles, and so forth. Above we have $\{ v_1 = 1, v_2 = 0, v_3 = 1, v_4 = 1, v_5 = 0, v_6 = 0, \dots \}$. Since no ball-number can be repeated in a cycle reduction, the cycle lengths must add up to the number N of balls.

$$v_1 + 2 v_2 + 3 v_3 + 4 v_4 + 5 v_5 + \dots + N v_N = N \tag{3.B.12}$$

So the number of different classes of permutations is equal to the number of *partitions* of the integer N .

For $N=2$ there are only two classes of two permutations.

$$\begin{aligned} &\text{Class } \{ v_1 = 2, v_2 = 0 \} \text{ corresponding to partition : } 2 = 1 + 1 \\ &\quad \text{One permutation : } (1)(2) \\ &\text{Class } \{ v_1 = 0, v_2 = 1 \} \text{ corresponding to partition : } 2 = 2 \\ &\quad \text{One permutation : } (12) \end{aligned} \tag{3.B.13}$$

For $N=3$ there are three classes of six permutations.

$$\begin{aligned} &\text{Class } \{ v_1 = 3, v_2 = 0, v_3 = 0 \} \text{ corresponding to partition : } 3 = 1 + 1 + 1 \\ &\quad \text{One permutation : } (1)(2)(3) \\ &\text{Class } \{ v_1 = 1, v_2 = 1, v_3 = 0 \} \text{ corresponding to partition : } 3 = 2 + 1 \\ &\quad \text{Three permutations : } (12)(3), (13)(2), (23)(1) \\ &\text{Class } \{ v_1 = 0, v_2 = 0, v_3 = 1 \} \text{ corresponding to partition : } 3 = 3 \\ &\quad \text{Two permutations : } (123), (132) \end{aligned} \tag{3.B.14}$$

The number of permutations in each partition class is given by a relatively simple combinatorial formula. To derive it one needs only consider the redundancy of the cycle labeling which was seen after (3.B.8), for example. Each M -cycle can be written M ways by cycling the numbers as shown in the tri-cycle in (3.B.8). If there are v_M such M -cycles in a permutation then there are M^{v_M} such reorderings that do not change the permutation at all. Also, since there are different numbers in each cycle they commute. So there are $v_M!$ reorderings of the v_M commuting cycles that give the same permutation, again. Dividing all these possibilities into $N!$ gives the number of distinct partition class numbers.

$$\begin{aligned} \text{Number in partition class } v_1 v_2 v_3 v_4 \dots &= \frac{N!}{v_1! 1^{v_1} v_2! 2^{v_2} v_3! 3^{v_3} v_4! 4^{v_4} \dots} \tag{3.B.15} \\ \text{where: } N &= v_1 + 2v_2 + 3v_3 + 4v_4 \dots \end{aligned}$$

Exercise: Classify and enumerate the permutations for $N=4$ and $N=5$.
(Check against (3.B.15))

Chapter 3 Problems

Mirror-Mirror (Who's the fairest eigenvector?)

3.1.1 Compute the eigenvectors, eigenvalues, spectral decomposition and d-trans matrix for each of the mirror operations (a) thru (d) in the *Mirror-Mirror* problem. Where possible, tell physical or geometric significance.

(a) Use spectral decomposition's to derive inverse $I=1/T$ and (all) square roots $X=\sqrt{T}$ such that $X^2=T$. (How many square roots does each have? Are any physically "do-able?")

(b) (extra-credit) Use c-d to invent a "slide rule" that correctly rotates U(2) electron and photon states.

Circle-Squash Switched

3.1.2 The discussion at the beginning of Sec. 1.6 showed that a unit circle is mapped onto an ellipse $\langle r|T^{-2}|r\rangle=1$ by matrix

$$T = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}. \text{ Consider the same mapping by "switched" matrix } S = \begin{pmatrix} 1/2 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

(a) Find eigenvalues of S and S^{-2} . Spectrally decompose S and plot its eigenvectors.

(b) Let $T^{-1}|r\rangle=|c\rangle=\langle r|T^{-1}$ or $T|c\rangle=|r\rangle=\langle c|T$ so $\langle r|T^{-1}|r\rangle=\langle c|r\rangle=\langle c|T|c\rangle$. Suppose all c-vectors lie on a curve $\langle c|T|c\rangle=1$ Discuss curve algebraically and plot this curve and the mapped $\langle r|T^{-1}|r\rangle=1$ curve.

(c) Let $S^{-1}|r\rangle=|c\rangle=\langle r|S^{-1}$ or $S|c\rangle=|r\rangle=\langle c|S$ so $\langle r|S^{-1}|r\rangle=\langle c|r\rangle=\langle c|S|c\rangle$. Suppose all c-vectors lie on a curve $\langle c|S|c\rangle=1$ Discuss curve algebraically and plot this curve and the mapped $\langle r|S^{-1}|r\rangle=1$ curve.

(d) *By conic geometry, derive a map $M|c\rangle=|r\rangle$ of any real vector $|c\rangle$ by real-symmetric matrix M.

Dagger Your Own Ket

3.1.3 Most quantum matrices have simple relations between eigenvalues ϵ_m and their conjugates ϵ_m^* , eigenbras $|e_m\rangle$ and kets $\langle e_m|$, projectors P_m and their †-conjugates $(P_m)^\dagger$, and diagonalizing (d-tran) transformations T and their inverses T^{-1} . Let's see what these relations are for...

(a) ...a Hermitian matrix $M = H$ such that $H = H^\dagger$ by spectrally decomposing and diagonalizing a general 2x2 reflection

matrix $H = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$. (Are its eigenvectors meaningful? Discuss.)

(b) ...a Unitary matrix $M = U$ such that $U^{-1} = U^\dagger$ by spectrally decomposing and diagonalizing a general 2x2 rotation

matrix $U = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$. (Are its eigenvectors meaningful? Discuss.)

(c) Find all the square-roots of H and of U. (Test them. There are more than two of each!)

Home on Lagrange

3.1.4 Functional spectral decomposition (3.1.17) is related to Lagrange functional interpolation (3.1.18). Use (3.1.18) to approximate $\sin x$ given only that $\sin 0=0$, $\sin \pi/2=1$, and $\sin \pi=0$. Compare your approximation to order-2 Taylor series approximation of $\sin x$ around $x=\pi/2$.

Bras-ackwards

3.1.5 See if you can work the spectral decomposition ideas backwards by doing the following "inverse" eigenvalue problems. (Hint: Use ket-bras and \otimes . Normalize first!)

(a) Find a Hermitian 3x3 matrix H that satisfies.

$$H \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad H \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad H \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 9 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

(b) Write down and test at least one square root \sqrt{H} . (How many square roots are there?)

Cures for Nilpotency

3.1.6 Can a *nilpotent* matrix \mathbf{N} ($\mathbf{N}^m = \mathbf{0}$, \mathbf{N}^{m-1} not zero, integer $m > 1$) be Hermitian $\mathbf{N} = \mathbf{N}^\dagger$

- (a) ...for $m=2$? , (b) ...for other m ? (Experiment with 2×2 matrices first.)
 (c) Use this and exercise *Dagger Your Own Ket* to prove Hermitian matrices must be diagonalizable.

Truly Secular

3.1.7 The coefficients a_k of the general $n \times n$ secular equation (3.1.5d) and (3.1.5f) of \mathbf{M} depend on matrix coefficients M_{ij} and on eigenvalues ϵ_m .

- (a) Do they depend on which basis you use to represent \mathbf{M} ? Why or why not?
 (b) For a general 4×4 matrix ($n=4$), compute functions $a_k = a_k(\epsilon_m)$ in an orderly way that clearly shows how they come out for general n .
 (c) For a general 4×4 matrix ($n=4$), compute functions $a_k = a_k(M_{ij})$ in an orderly way that clearly shows how they come out for general n . Use the ϵ -expansion in Appendix 1.A and (b) above to help express answer in terms of diagonal minor determinants. (NOTE: This is a "crucial" problem whose solutions belongs in your lab "journal" or equivalent.) May do successively $n=2, 3$, until a pattern emerges.

Adjunct Junk

3.1.8 Given (1.A.5) or $\mathbf{A} \mathbf{A}^{ADJ} = \mathbf{1}$ ($\det|\mathbf{A}|$) with $\mathbf{A} = \mathbf{M} - \lambda \mathbf{1}$ show that \mathbf{A}^{ADJ} has \mathbf{M} eigenkets $|\lambda\rangle$ if λ is an eigenvalue of \mathbf{M} .

Does \mathbf{A}^{ADJ} also harbor \mathbf{M} 's eigenbras? Use $\mathbf{M} = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$ as an example.

Pair 'em up

3.1.9 An $n \times n$ pairing matrix Π has 1 for all n^2 matrix elements $\Pi_{ij} = 1$. It's used in superconductivity theory and nuclear structure.

- (a) Use 1(c) above to help derive its eigenvalues and spectral decomposition. (Or, you may develop the theory by doing successively $n=2, 3$, until the pattern emerges.)
 (b) Does the matrix $\Pi + (\text{const.})\mathbf{1}$ have the same eigenvectors? eigenvalues? as Π . Explain.

All Together Now

3.1.10 Show how to do a simultaneous spectral decomposition using the projector splitting technique.. (a) Spectrally

decompose $\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, and $\mathbf{B} = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$.

- (b) Calculate the "ridiculous function" $\mathbf{B}^{\mathbf{A}}$ of these two matrices.

A Perturbing Problem

3.2.1 Find eigenvalues (to $\pm 1\%$) of matrix $\mathbf{M} = \begin{pmatrix} 2 & 0.1 & 0.3 \\ 0.1 & 3 & 0.2 \\ 0.3 & 0.2 & 4 \end{pmatrix}$ using perturbation theory.

A Permuted Problem

3.2.2 (a) As in Appendix 3.B show cycle structure of all permutations in symmetric group S_4 and S_5 .

- (b) Write permutation $(\mathbf{p})|12345678\rangle = |25386741\rangle$ in cycles. How many (\mathbf{p}) in its S_8 class?

Unit. 1 Review Topics and Formulas

Transformation Matrix

General $n \times n$ T Matrix

$$\text{Bras: } \begin{pmatrix} \langle 1| \\ \langle 2| \\ \langle 3| \\ \vdots \end{pmatrix} \begin{pmatrix} \langle 1|1'\rangle & \langle 1|2'\rangle & \langle 1|3'\rangle & \dots \\ \langle 2|1'\rangle & \langle 2|2'\rangle & \langle 2|3'\rangle & \dots \\ \langle 3|1'\rangle & \langle 3|2'\rangle & \langle 3|3'\rangle & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{matrix} |1'\rangle \\ |2'\rangle \\ |3'\rangle \\ \dots \end{matrix}$$

Transformation Matrix

1. Photon Polarization

$$\begin{pmatrix} \langle x|x'\rangle & \langle x|y'\rangle \\ \langle y|x'\rangle & \langle y|y'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Transformation Matrix

2. Electron Polarization

$$\begin{pmatrix} \langle \uparrow|\uparrow'\rangle & \langle \uparrow|\downarrow'\rangle \\ \langle \downarrow|\uparrow'\rangle & \langle \downarrow|\downarrow'\rangle \end{pmatrix} = \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}$$

Axiom 1: The absolute square $|\langle j|k'\rangle|^2 = \langle j|k'\rangle^* \langle j|k'\rangle$ gives the probability for state- j of a system in state- $k'=1'$ to n' from one sorter and then forced to choose between states $j=1$ to n by another sorter.

Axiom 2: The complex conjugate of an amplitude gives its reverse: $\langle j|k'\rangle^* = \langle k'|j\rangle$

Axiom 3: If identical analyzers are used twice or more the amplitude for a passed state- k is one,

$$\text{and for all others it is forever zero: } \langle j|k\rangle = \delta_{jk} = \begin{cases} 1 \text{ if: } j=k \\ 0 \text{ if: } j \neq k \end{cases} = \langle j'|k'\rangle \text{ (ORTHONORMALITY)}$$

Axiom 4. Ideal sorting followed by ideal recombination of amplitudes has no effect:

$$\langle j''|m'\rangle = \sum_{k=1}^n \langle j''|k\rangle \langle k|m'\rangle \Rightarrow \mathbf{1} = \sum_{k=1}^n |k\rangle \langle k| = \sum_{k=1}^n \mathbf{P}_k \text{ (COMPLETENESS)}$$

The secular equation $\det|\mathbf{M} - \epsilon\mathbf{1}| = 0 = (-1)^n (\epsilon^n + a_1\epsilon^{n-1} + a_2\epsilon^{n-2} + \dots + a_{n-1}\epsilon + a_n)$ where:

$$a_1 = -\text{Trace}\mathbf{M}, \dots, a_k = (-1)^k \sum \text{diagonal } k\text{-by-}k \text{ minors of } \mathbf{M}, \dots, a_n = (-1)^n \det|\mathbf{M}|$$

The Hamilton-Cayley (HC) equation $\mathbf{0} = (\mathbf{M} - \epsilon_1\mathbf{1})(\mathbf{M} - \epsilon_2\mathbf{1}) \dots (\mathbf{M} - \epsilon_n\mathbf{1})$

Projection operators: $\mathbf{P}_k = \frac{\prod_{j \neq k} (\mathbf{M} - \epsilon_j\mathbf{1})}{\prod_{j \neq k} (\epsilon_k - \epsilon_j)}$ are eigenoperators for \mathbf{M} such that: $\mathbf{P}_k \mathbf{M} = \mathbf{M} \mathbf{P}_k = \epsilon_k \mathbf{P}_k$

\mathbf{P}_k satisfy projector orthonormality $\mathbf{P}_j \mathbf{P}_k = \delta_{jk} \mathbf{P}_k$ and projector completeness $\mathbf{1} = \mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_n$

and: spectral decompositions of an operator \mathbf{M} . $\mathbf{M} = \epsilon_1 \mathbf{P}_1 + \epsilon_2 \mathbf{P}_2 + \dots + \epsilon_n \mathbf{P}_n$

$$f(\mathbf{M}) = f(\epsilon_1) \mathbf{P}_1 + f(\epsilon_2) \mathbf{P}_2 + \dots + f(\epsilon_n) \mathbf{P}_n$$

The old " $\mathbf{1}=\mathbf{1}\cdot\mathbf{1}$ trick" will be used later in symmetry analysis and spectroscopic theory

$$\mathbf{1} = \mathbf{1} \cdot \mathbf{1} = (\mathbf{P}_1^G + \mathbf{P}_{-1}^G)(\mathbf{P}_2^H + \mathbf{P}_{-2}^H) = \mathbf{1} = (\mathbf{P}_1^G \mathbf{P}_2^H + \mathbf{P}_1^G \mathbf{P}_{-2}^H + \mathbf{P}_{-1}^G \mathbf{P}_2^H + \mathbf{P}_{-1}^G \mathbf{P}_{-2}^H)$$

Perturbation expansion for eigenvalue nearest $E_1 = H_{11}$: (For order higher than 2: Caution and good luck!)

$$\lambda = E_1 + \sum_{j \neq 1}^N \frac{H_{1j} H_{j1}}{(E_1 - E_j)} + \sum_{j \neq 1}^N \sum_{k \neq 1, j}^N \frac{H_{1j} H_{jk} H_{k1}}{(E_1 - E_j)(E_1 - E_k)} + \sum_{j=1}^N \sum_{k \neq j}^N \sum_{\ell \neq j, k}^N \frac{H_{1j} H_{jk} H_{k\ell} H_{\ell 1}}{(E_1 - E_j)(E_1 - E_k)(E_1 - E_\ell)} \dots$$

Unit 1 Exam **Quantum Mechanics 5413** Dirac Notation and Matrix Algebra

(40 pts)

1. We're given the following base state definitions of transformation operator **R**.

$$|1'\rangle = \mathbf{R}|1\rangle = |2\rangle ,$$

$$|2'\rangle = \mathbf{R}|2\rangle = |3\rangle ,$$

$$|3'\rangle = \mathbf{R}|3\rangle = |4\rangle ,$$

$$|4'\rangle = \mathbf{R}|4\rangle = |1\rangle ,$$

- (a) Write down a matrix representation for **R** in the $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}$ basis in Dirac notation and numerically.
- (b) Use (a) to compute a representation of **R**², **R**³, and **R**⁴, too.
- (c) Write down a matrix representation for **R** in the $\{|1'\rangle, |2'\rangle, |3'\rangle, |4'\rangle\}$ basis in Dirac notation and numerically. Is it different from the result in (a)? Why or why not?
- (d) By examining powers **R**^p deduce the Hamilton Cayley equation and secular equation of **R** matrix.
- (e) Write down minimal equation and eigenvalues for **R**.
- (f) Can matrix **R** be spectrally decomposed and diagonalized? How do you tell?
- (g) Can all the matrices **R**, **R**², **R**³ ..., **R**^p be simultaneously decomposed by a single set of projectors and transformation matrix. How do you know?
- (h) If (f) is "Yes" do spectral decomposition and diagonalization of **R**.
- (i) If (g) is "Yes" do spectral decomposition and diagonalization of **R**^p for any power *p*.

(30 pts)

2. The results from the preceding problem may help to spectrally decompose the following general type of matrix. If so, explain why and use the results to find its eigenvectors and eigenvalues in terms of constant parameters *A*, *B*, *C* and *D*. If not, explain why not.

$$\mathbf{H} = \begin{pmatrix} D & A & B & C \\ C & D & A & B \\ B & C & D & A \\ A & B & C & D \end{pmatrix} \text{ Sketch levels for case } A=B=C=0.2 \text{ and } D=1.$$

(20 pts)

3. (a) Write a 2nd order perturbation expression for the eigenvalues of

$$\mathbf{H} = \begin{pmatrix} D_1 & A & B & C \\ C & D_2 & A & B \\ B & C & D_3 & A \\ A & B & C & D_4 \end{pmatrix}$$

in terms of parameters *A*, *B*, and *C* for *D*₁ = 1, *D*₂ = 2, *D*₃ = 3, *D*₄ = 4.

- (b) Sketch levels for case *A*=*B*=*C*=0.2.
- (c) Can you use your expression (a) on the matrix in problem 2? Why or why not? Explain while giving a brief discussion of the requirements for a valid perturbation result.

(10 pts)

4. The transformations **R**, **R**², **R**³, in Problem 1 also behave like *permutations*. How?

- (a) Give the cycle structure and notation for each. How many distinct **R**^p (*p* integral) exist ?
- (b) How many different permutations can you make by considering all possible arrangements of numbers 1,2,3, and 4 in the cycles of **R**? Classify them by cycles.

