Spin

Michael Weiss

The Facts I will begin by stating the bare facts briefly, at the cost of some precision.

Any quantum-mechanical system possesses angular momentum, sometimes called *spin*[1]. Angular momentum is quantized, that is it is always an integer multiple of $\hbar/2$. (Here, $\hbar = h/2\pi$, where h of course is Planck's constant.) Traditionally, j stands for angular momentum[2]. If we use so-called natural units of measurement (and we will), then $\hbar = 1$ and possible values of j are $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$.

Say we measure the component of angular momentum along some axis, say the z-axis. Let j_z stand for this component. [3] Then j_z is also quantized, and possible values for j_z are

$$-j, -j+1, \ldots, j-1, j$$

A more precise definition of j is the maximum possible value for j_z. For example, a system with angular momentum one-half (called a "spin-$\frac{1}{2}$ system" for short) will have $j_z = \pm\frac{1}{2}$; a system with angular momentum 1 (spin 1) will have $j_z = -1, 0$, or $+1$. The same result holds for the component of angular momentum measured along *any* axis.

The total magnitude of angular momentum is given by $\sqrt{j_x^2 + j_y^2 + j_z^2}$; if this is measured somehow[4], the result will be $\sqrt{j(j+1)}$. In classical mechanics, we would get j and not $\sqrt{j(j+1)}$ (as we will see in a moment).

Suppose we combine two systems together, one with angular momentum j, the other with angular momentum j'. The resulting composite system will have

[1] Historically, the term *spin* had (and sometimes has) a narrower meaning than *angular momentum*. Picture an electron in an atom. Sacrificing some accuracy on the altar of visual imagery, think of the electron as orbiting about the nucleus, and spinning on its axis. The orbital motion gives rise to so-called orbital angular momentum; the spinning gives rise to so-called intrinsic or spin angular momentum.

[2] And l traditionally stands for orbital angular momentum, and s stands for spin angular momentum.

[3] People sometimes use m_z instead of j_z to denote this.

[4] Not so trivial, because the three component measurements are mutually incompatible; it is theoretically impossible to measure j_x and j_y simultaneously (for example). Nonetheless, the combination $j_x^2 + j_y^2 + j_z^2$ *is* measurable.

angular momentum equal to one of the values

$$|j-j'|, |j-j'|+1, \ldots, j+j'$$

For example, combining two spin-$\frac{1}{2}$ systems can result in a spin-0 system or a spin-1 system.

This addition rule for angular momenta lies behind a wealth of physical phenomena. A simple example: suppose an atom absorbs a photon. A photon always has spin 1. Suppose the atom starts out in a state with $j = 1$. After absorbing the photon, the atom must make a transition to a state with $j = 0, 1$, or 2. The same conclusion holds when an atom emits a photon. This three-way choice ultimately manifests itself as a triplet of spectral lines.

Now let's be a little more precise. In classical mechanics, angular momentum is a vector, say $\mathbf{j}$, with components j_x, j_y, j_z in some coordinate system, and magnitude $j = |\mathbf{j}| = \sqrt{j_x^2 + j_y^2 + j_z^2}$. Any component, say j_z, can range in value from $-j$ to $+j$.

In quantum mechanics, we have a (complex) Hilbert space of state-vectors (say H) for any system. A state-vector $v \in H$ specifies a state of the system, and v and w specify the same state if and only if $v = cw$ for some non-zero complex number c. Let Q be some classical, real-valued variable that you can measure (like energy or momentum). The "quantum version" of Q is a Hermitian operator on H. The eigenvalues of this operator are the possible values you can get from measuring Q. The quantum system has a definite value for Q if and only if the system is in an eigenstate of the operator. If not, the act of measurement will serve to cast the system into such an eigenstate, with probabilities that can be computed by the rules of quantum mechanics.

Applying this prescription to angular momentum, we see that j_x, j_y, j_z all must be Hermitian operators. It turns out now that the Hilbert space of quantum states decomposes, in the most general case, into a direct sum:

$$H = H_0 \oplus H_{1/2} \oplus \ldots \oplus H_j \oplus \ldots$$

where each summand in turn decomposes:

$$H_j = \sum_{m=-j}^{j} K_{jm}$$

and each element of K_{jm} is an eigenvector of j_z with eigenvalue m. (Here, m ranges in steps of 1 from $-j$ to $+j$.) In particular cases, some of these summands may be missing. For example, if the quantum system has definite angular momentum j, then $H = H_j$.

Each K_{jm} is of course invariant under j_z, i.e., we have an invariant direct-sum decomposition of H for the operator j_z. The K_{jm} are *not* invariant under j_x or

j_y, but it turns out that the H_j *are* invariant under all three operators j_x, j_y, and j_z. Put another way, the invariant direct-sum decompositions for j_x and j_y have the same H_j's but different K_{jm}'s.

Finally, the operator $j_x^2 + j_y^2 + j_z^2$ is itself a Hermitian operator; each H_j is an eigenspace, with eigenvalue $j(j+1)$.

This excursion into Hilbert spaces should make mathematically-minded folk more comfortable with the catalog of "spin facts". The space H_j is just the space of states that have "spin j"; the subspace K_{jm} is the space with component m along the z-axis. I haven't discussed the combination rules; this translates into statements about the direct-sum decomposition of tensor products. Nor have I explained the curious appearence of the term $j(j+1)$; this is bound up with the non-commutativity of j_x, j_y, j_z.

Historically, the rules for spin came from playing with experimental data. The rules worked but remained mysterious. The birth of quantum mechanics came later, at the hands of Heisenberg, Schrödinger, and Dirac. Spin fell into place soon after that.

Quantum mechanics is mysterious, of course, as numerous philosophical treatises attest. But during the crucial period from Bohr's first great work (in 1913) to Heisenberg's discovery of matrix mechanics (in 1925), certain technical facts sowed confusion above and beyond the general quantum "spookiness". I will single out one of these for special attention: the spin of the electron. In 1926, Goudsmit and Uhlenbeck proposed (correctly) that the electron has spin one-half. Only integer spin quantum systems have classical counterparts, as we will see. The fractional spin of the electron lurked amid the general confusion during the heyday of the "old quantum theory" (1913–1925), bedevilling physicists. (Goudsmit and Uhlenbeck actually made their proposal in the language of the old quantum theory, not the newly minted quantum mechanics. The next year Pauli showed how to incorporate spin into quantum mechanics.)

In these notes, I will roam through the history of the old quantum theory, zeroing in on the adumbrations of modern quantum spin. (I will intersperse remarks on what the pioneers were "really doing".) I will not lay out the whole array of modern mathematical toys that makes Everything Clear. The reader whose appetite has been whetted must turn to standard textbooks for that. But I will try paint an impressionistic picture of the luminous synthesis, with a few broad brush strokes and one or two detailed pointillistic patches.

Entr'act In 1900, Planck's paper on black-body radiation introduced his constant h; in 1905, Einstein's paper on the photo-electric effect introduced light-quanta (later called photons). I hasten past these developments with just a smattering of comments. Both phenomena dealt with the interaction of light and matter. Late nineteenth century physicists felt they understood light, thanks to Maxwell. They knew less about matter, and knew they knew less, but had no reason to suspect that Newton's laws didn't hold "all the way down". But the interaction between the light and matter lay at the research frontier.

Einstein proposed that light of frequency ν was composed of quanta, each with energy $h\nu$. This flies in the face of Maxwell's identification of light with electromagnetic waves. Moreover, Einstein's relation weds a particle concept (the energy of one photon) to a wave concept (the frequency). (Einstein regarded his hypothesis as provisional— a guide to a more complete theory.)

The Old Quantum Theory Bohr's 1913 model of the hydrogen atom really starts the story. Nineteenth and early twentieth century experimentalists heaped up data on the spectra of atoms. Besides the plain old spectra of hydrogen, helium, sodium, etc., these physicists studied ionized atoms, atoms in electric fields (the Stark effect), atoms in magnetic fields (the Zeeman effect), atoms in crossed electric and magnetic fields; they discovered the so-called "fine-structure" of spectra, where one spectral line under higher resolution splits into two or more spectral lines; and so on, and so on, for thousands of journal pages.

Bohr's model solidified regularities in this data already partly noted by earlier workers (notably Balmer and Ritz). Bohr's transition picture (as I will call it) states that an atom has a discrete set of energy levels. When the atom emits a photon, it loses energy, changing from (say) energy level E_i to level E_j, so the photon has energy $E_i - E_j$ and hence frequency

$$\nu = (E_i - E_j)/h$$

(This is known as the Bohr frequency condition.) So to explain the spectrum of an atom, we must identify all its energy levels. Of course, this picture depends fundamentally on the Einstein relation. It is utterly incompatible with any classical picture in which an orbiting electron gradually changes from one orbit to another, emitting radiation as it decellerates. Indeed, classically there should be no stable orbits at all!— as noted in all histories of quantum theory.

The transition picture survives in modern quantum mechanics, although naturally much refined. The electromagnetic field is a quantum system; so is the atom; the two are coupled to form one combined quantum system. The interaction between the atom's charged particles and the electromagnetic field shows up as a term in the Hamiltonian (the energy operator, which governs the time evolution of the system). If this so-called coupling term is neglected, then the mathematics predicts that the atom should have discrete stable energy levels (i.e., if the atom is in an eigenstate for a particular energy level, then it will stay in that eigenstate forever.)

Just as the energy of the atom is quantized, so too is the energy of the electromagnetic field at a particular frequency. This falls out from computations with the quantum version of Maxwell's equations. We interpret this result as saying that electromagnetic field consists of quanta (or photons). Einstein's relation $E = h\nu$ can be deduced, rather than postulated.

Finally, if the coupling term between the field and the charged particles is *not* neglected, transitions become possible between formerly stable eigenstates. Both the atom and the field change their state together. Bohr's frequency condition can be derived, placing a satisfying capstone on the theoretical development.

Returning to history's tangled tale... Bohr's 1913 paper did more than introduce the transition picture. He also derived the energy levels for the hydrogen atom, by a remarkably simple argument. He assumed that the orbital angular momentum of the atom's solitary electron was $n\hbar$, with n a positive integer. He assumed circular orbits, and applied Newton's laws, and the formula for the energy-levels of the atom popped right out. And it was right! (Apart from second-order corrections[5].)

Once Bohr opened the gates, other quantum soldiers flocked in to consolidate the victory. Sommerfeld (in 1916) extended Bohr's model to allow for elliptical orbits. He included relativistic corrections, and calculated the effects of magnetic fields. Sommerfeld assembled a cadre of students, in Munich. Bohr had his group at Copenhagen, and Born started one at Göttingen. Heisenberg and Pauli began their careers as students in Sommerfeld's seminar on spectra, and studied at all three centers.

Almost all work during the period 1913–1925 shared a common approach, now called the *old quantum theory.* One began with a classical treatment of a dynamical problem— for example, elliptical orbits. One slapped *quantum conditions* on top. Quantum conditions stated that certain classical quantities had to be integer multiples of $\hbar$— like the orbital angular momentum of the electron, in Bohr's theory of the hydrogen atom. (Sommerfeld added other quantum conditions.) Quantum conditions plus classical physics yielded a discrete set of energy levels, which in turn yielded spectra via Bohr's frequency condition.

These forays met initially with great success, but ultimately with crushing defeat. Generally speaking, the Bohr-Sommerfeld approach handled two-body problems reasonably well, but collapsed when faced with three-or-more-body problems. The hydrogen atom is a two-body problem: an electron orbits a proton. Bohr himself extended the theory to "hydrogenic" atoms— atoms that consist of a core of tightly-bound electrons around the nucleus, and a single loosely-bound electron orbiting further out. Hydrogenic atoms are "approximately" two-body systems: core-plus-electron. As soon as the theoreticians turned to three-body problems, theory and experiment parted company. Helium (two electrons and a nucleus) and the singly-ionized hydrogen molecule

[5]Specifically, Bohr's energy levels agree exactly with those derived from Schrödinger's equation, when we neglect the spin of the electron, the spin of the proton, and relativistic effects. (Spin itself is sometimes considered a relativistic effect.)

(one electron and two nuclei) spelled doom for the old quantum theory.

The Hydrogen Atom, Then and Now Let us look at a simple model of hydrogen atom: we neglect the spin of the proton and the electron, and relativistic effects. What remains is a point charge in an inverse-square force field— the classical Kepler problem.

With these simplifications, the states of the hydrogen atom can be specified by giving three integer labels (n, l, m), with:

$$\begin{aligned} n &= 1, 2, \ldots \\ l &= 0, 1, \ldots, n-1 \\ m &= -l, -(l-1), \ldots, (l-1), l \end{aligned}$$

The labels (n, l, m) are called *quantum numbers*, and made their debut in Sommerfeld's elliptical orbit model. They correspond to certain classical features of the orbit:

$$\begin{aligned} n &\leftrightarrow \sqrt{\text{semimajor axis}} \\ l &\leftrightarrow \text{orbital angular momentum} \\ m &\leftrightarrow \text{orientation of the orbit} \end{aligned}$$

Sommerfeld's quantum conditions stated that these three classical quantities were restricted to integer values (and in fact the collection of values given above).[6] I should really say "integer multiples of $\hbar$", but from now on I will adopt so-called natural units of measurement, which are chosen so that $\hbar = 1$.

I should make one refinement to this description of the "old quantum" viewpoint. The quantum conditions apply only to so-called *stationary* orbits. Bohr offered no details for what happened during a transition from one stationary orbit to another (a "quantum jump"), nor could he explain why these orbits were stationary. Initially, physicists probably regarded these questions as topics for future research.

In the post-1925 reformulation, we have to solve a particular instance of Schrödinger's equation, subject to certain boundary conditions. The space of all solutions is a Hilbert space, with a basis $\{v_{nlm}\}$. The vector v_{nlm} is simultaneously an eigenvector of three operators:

E, the energy
L, the magnitude of the orbital angular momentum

[6]Sommerfeld gave a single more general quantum condition which implied the quantum conditions for n, l, and m as special cases. So Sommerfeld's approach was not so ad-hoc as this summary makes it look.

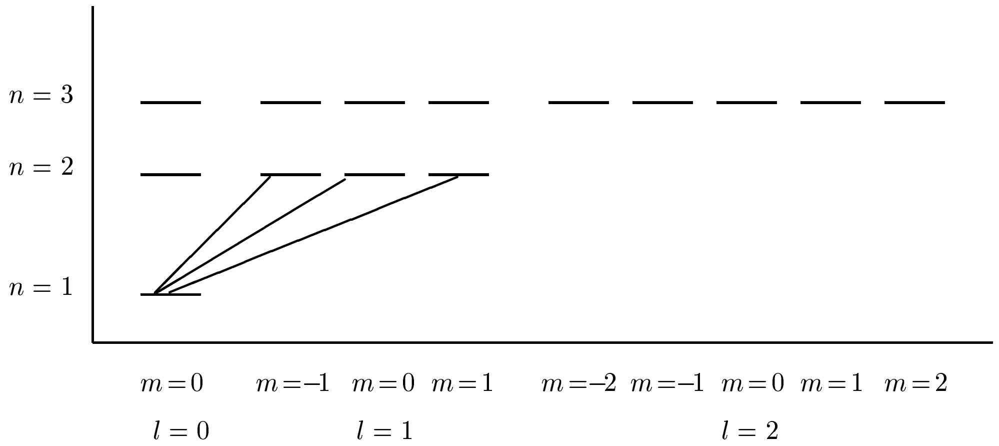

Figure 1: Term Scheme for Hydrogen

L_z, the z-component of the orbital angular momentum

So we can say that if the hydrogen atom is in state v_{nlm}, then it has a definite energy, and its angular momentum has a definite magnitude and z-component. In fact, the eigenvalues for v_{nlm} are:

$$\begin{aligned} Ev_{nlm} &= \frac{R}{n^2} v_{nlm} \\ Lv_{nlm} &= \sqrt{l(l+1)}\, v_{nlm} \\ L_z v_{nlm} &= m v_{nlm} \end{aligned}$$

where R is a physical constant known as Rydberg's constant.

The modern equivalent of the old notion of "stationary orbit" is "eigenstate of the energy operator". Such eigenstates do not change with time, and they possess a definite value for the energy. (These facts are closely related.) Schrödinger's equation, in fact, amounts to $Ev = \lambda v$, where E is the energy operator.

Figure 1 gives a pictorial representation of the basis $\{v_{nlm}\}$ in a form known as a *term scheme.* The horizontal lines stand for basis vectors (or equivalently, stationary quantum states); height gives energy, and transitions are indicated by slanted lines. (Only three transitions are pictured, to avoid clutter.)

From this simple diagram, many treasures flow. The next few paragraphs give a taste.

Degeneracy We see that all energy levels, except for $n = 1$ (the so-called ground state), contain several states. Physicists say that the energy levels are *degenerate*; in particular, the degree of degeneracy of level n is $1+3+5+\ldots+(2n-1)=n^2$. Mathematically, the n-th eigenspace of the energy operator has dimension n^2.

Moralists may abhor degeneracy, but physicists love it, for it adds spice and complexity. (Of course, the same may hold for moral degeneracy!) Multi-dimensional eigenspaces furnish room for interesting transformation groups. Physical applications of group representation theory really took off after the discovery of quantum mechanics.

Pertubations and Fine Structure Bohr's formula for the energy levels is quite simple:

$$E_{nlm} = \frac{R}{n^2}$$

This simplicity stems from the simplicity of the formula he used for the energy:

$$E = \frac{p^2}{2m_e} + \frac{\text{constant}}{r}$$

where p^2 is the square of the momentum, r is the radial distance from the proton to the electron, and m_e is the mass of the electron. The first term represents the kinetic energy, the second the potential energy due to the inverse square force (Coulomb potential). The quantum theorists came along later and said that E, p^2, and r are all really operators, and the energy levels are really eigenvalues, but the source of the simplicity remains the same. "Simplicity" is too vague a word: "symmetry" serves better, as we will see.

For example, the formula for E is spatially symmetric under rotations about the proton (because p^2 and r are unchanged by such rotations). Thus the formula for E_{nlm} cannot depend on m, which specifies the orientation of the orbit. If we impose a magnetic field along the z-axis, say, then E acquires a new term (of the form BL_z, where B is the magnetic field strength, and L_z is the z-component of orbital angular momentum). The new formula for E_{nlm} now depends on m, since L_z is not symmetric under arbitrary rotations. The magnetic field is referred to as a perturbation.

We can see the effects of the magnetic field in changes to the term scheme: some states move up in energy, some move down. Reduced symmetry has lead to reduced degeneracy. Transitions that formerly had the same ΔE now have slightly different ΔE's (see figure 2). This means that a single spectral line with no magnetic field will split into multiple spectral lines when the field is turned on. This is the famous *Zeeman effect*.

Another example: in 1916, Sommerfeld replaced Bohr's simple formula for E with its relativistic equivalent. This lead to a new formula for E_{nlm}, depending

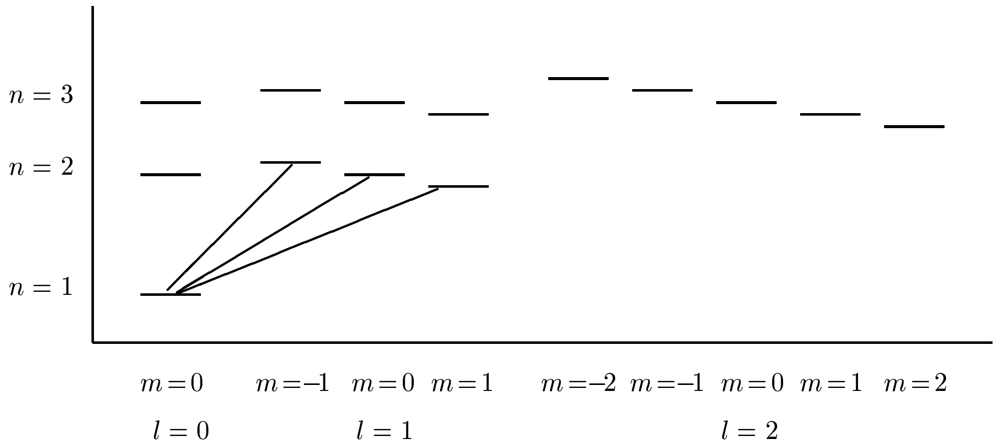

Figure 2: Term Scheme for Hydrogen in a (Strong) Magnetic Field

on both n and l (but not on m, of course). These relativistic corrections (or perturbations) account for the so-called *fine structure* of the hydrogen spectrum, known already to spectroscopists long before 1916.

A historical footnote: one modern author has written:

> When Dirac developed relativistic quantum mechanics, the relativistic Coulomb problem proved to be *exactly solvable* ... But the resulting formula for the energy levels was truly a surprise: *The new answer was precisely the old Sommerfeld formula!*
>
> How could this possibly be? Clearly Sommerfeld's methods were heuristic (Bohr quantization rules), out-dated by *two* revolutions (Heisenberg-Schrödinger nonrelativistic quantum mechanics and Dirac's relativistic quantum mechanics) and his methods obviously had no place at all for the electron spin ... So Sommerfeld's correct answer could only be a lucky accident, a sort of cosmic joke at the expense of serious minded physicists.

The author calls this the "Sommerfeld puzzle", and resolves it, but I will discuss it no further.

The Periodic Table From 1920 to 1923, Bohr applied his ideas to explain features of the periodic table; Bohr's work in this vein was corrected and extended by Pauli in 1924 (with an assist from Stoner).

Suppose we treat the states in the term scheme for hydrogen as slots to hold electrons. Let us hypothesize that each slot can hold at most two electrons. We start with hydrogen, and keep adding electrons, adding new electrons always to the lowest slot that has a vacancy. (Bohr christened this the *Aufbauprincip*, the building-up principle.) The energy levels of the slots may shift somewhat as a result of interactions between the electrons. This simple model (extremely naive from the modern viewpoint) serves to explain a remarkable number of chemical regularities. For example, energy level n can hold $2n^2$ electrons. We therefore get completely filled energy levels for 2 electrons, 2+8 electrons, 2+8+18 electrons, etc.— which correspond exactly to the noble gases.

Bohr was able to explain some properties of the rare earth elements using his version of this model[7]; he even predicted correctly that element 72 (not yet discovered) would *not* be a rare earth, but instead would resemble zirconium (contrary to what some chemists expected). Students of Bohr then discovered element 72 in zirconium ore samples, and element 72 was named hafnium in honor of the Latin name for Copenhagen.

Why two electrons per slot? Pauli proposed adding a new quantum number to the triple (n, l, m); nowadays we use s for this number, and recognize that it stands for the spin of the electron. So s is restricted to the values $\pm\frac{1}{2}$, and each state (or slot) in the term scheme above is really two states. In other words, our Hilbert space must be enlarged to a space with basis $\{v_{nlms}\}$. Pauli also proposed his famous *exclusion principle*: a state can hold at most one electron.

A historical note: Bohr made no detailed orbit calculations for multi-electron atoms. He did use some general intuitive principles stemming from a classical picture, e.g., a electron in a state of large n will be far away from the nucleus, and electrons with smaller n will partly "screen" the charge of the nucleus from the far-off electron. Nevertheless, the Bohr-Sommerfeld hydrogen model had a comforting, semi-classical feel to it: we compute the classical orbits, then decree that only some are permitted. For multi-electron atoms, Bohr dropped this connection. The term scheme becomes almost a combinatorial device.

Physicists (such as Born and Heisenberg) who did perform detailed orbit calculations found utter disagreement between theory and experiment— even for helium, a mere two electrons. No surprise, since the classical picture leaves out spin, the exclusion principle, and other purely quantum interactions between the electrons. Bohr intuited just how far to push the classical picture; he picked those approximation schemes that survive (reinterpreted) in quantum mechanics.

The Zeeman Effect Place an atom in a magnetic field, and (as Zeeman discovered) its spectral lines will split. Pais's book *Inward Bound* gives fascinating details of the experimental history.

[7]Bohr did not make use of the quantum number m; instead, he had n slots at level n ($l = 0, 1, \ldots, n-1$), each capable of holding $2n$ electrons. Note that this is actually incorrect.

For some atoms, spectral lines split in three under a magnetic field. This is known as the *normal Zeeman effect.* For other atoms, the spectrum displays a more complex pattern of splittings, known as the *anomalous Zeeman effect.* When the magnetic field becomes strong enough, however, some lines merge back together, and the anomalous Zeeman splitting coalesces smoothly into a normal splitting— this is called the *Paschen-Back effect.* Young Heisenberg wrestled with the anomalous Zeeman effect, as a student in Sommerfeld's seminar on atomic spectra (see Cassidy's biography *Uncertainty.*)

The full explanation of all this phenomena presented quite a puzzle to the quantum pioneers. Even today, fitting all the pieces together can be confusing. In the next few sections I will try to assemble this puzzle.

I treat the Zeeman effect in such gory detail for two reasons. First, historical interest: the Zeeman effect gave the first hints of the complexities of quantum spin. Many of the strange "spin facts" I stated earlier were first discovered empirically, from poring over spectra and noticing patterns.

Second, "gestalt": the very complexity of the Zeeman effect means that any explanation of it must bring together several aspects of quantum mechanics.

We can get started by comparing figures 1 and 2. We see that spectral lines will split if energy levels move up or down (i.e., if the energy levels of quantum states change). To progress further, one needs a formula for the energy levels of the states of an atom— or rather, a formula for the change in energy level due to the magnetic field.

Say we write $E = E^0 + E^{\mathrm{mag}}$, where E^0 is the energy with no magnetic field, and E^{mag} is the energy due to the magnetic field. (I will add quantum-number subscripts (like we had with E_{nlm}) in a moment.) Here is a formula for E^{mag} that accounts for the normal Zeeman effect:

$$E^{\mathrm{mag}}_{LM_LSM_S} = \mathrm{constant}\, B(M_L + 2M_S) \tag{1}$$

where L, M_L, S, and M_S are quantum numbers that I will discuss below, and B is the strength of the magnetic field. Here is a formula that accounts for the anomalous Zeeman effect:

$$E^{\mathrm{mag}}_{JLSM_J} = \mathrm{constant}\, BM_J \left(1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}\right) \tag{2}$$

where J and M_J are another pair of quantum numbers to be discussed below. The complicated fraction built out of J, L, and S on the right hand side is known as the Landé g-factor.

So there are four pieces to the puzzle:

1. What is the meaning of the six quantum numbers L, S, J, M_L, M_S, and M_J?

2. How do the "normal" and "anomalous" formulas (formulas 1 and 2) account for the details of the Zeeman splitting?

3. Where do the normal and anomalous Zeeman formulas come from?

4. How does the anomalous Zeeman formula turn into the normal Zeeman formula in the Paschen-Back limit of strong magnetic fields?

The next four sections tackle each of these questions in turn.

Angular Momentum Quantum Numbers *What is the meaning of the six quantum numbers L, S, J, M_L, M_S, and M_J?*

The "old quantum", Bohrish-Sommerfeldian notion of quantum numbers ran like this:

> Certain classical quantities are constrained, for a stationary quantum state, to take on only integer values. These values are called the quantum numbers of the state.

Generalizing this just a bit, we allow any discrete set of real numbers in place of the integers (e.g., the set $\{\sqrt{j(j+1)}\}$ as j ranges over the integers; or more significantly, as j ranges over the set of integers and half-integers).

Dressing this up in mathematical language, we have a space (actually a differential manifold) Σ that represents the set of all classical states. Classical quantities are continuous real-valued functions on Σ. Bohr and Sommerfeld assumed that nature really permits only a subset Σ_0 to occur as stationary states. They assumed further, for certain classical $f : \Sigma \to \mathbf{R}$, that the image $f(\Sigma_0)$ was a discrete subset of $\mathbf{R}$.

The "new quantum" notion of quantum numbers (à la Heisenberg, Schrödinger, and Dirac) runs like this:

> Classical quantities correspond to Hermitian operators on a complex Hilbert space. Quantum states are specified by elements of the space. If a Hermitian operator Q has a discrete spectrum, and v is an eigenvector of Q with eigenvalue q, then we say that q is a quantum number of the state specified by v.

I will start with the "old quantum" explanation of our six quantum numbers. (The "new quantum" version will have to wait till we discuss the Paschen-Back effect.) Consider first the hydrogen atom. The electron possesses orbital angular momentum, given by a vector[8] $\mathbf{l}$. Since the electron is spinning, it has also spin

[8] We mean an ordinary 3-space vector, not a vector in a (complex) Hilbert space.

angular momentum, given by a vector $\mathbf{s}$. The total angular momentum of the system[9] is then given by the vector $\mathbf{j} = \mathbf{l} + \mathbf{s}$.

For a multi-electron atom, we let $\mathbf{L}$ be the sum of the $\mathbf{l}$'s over all the electrons; $\mathbf{S}$ is likewise the sum of all the $\mathbf{s}$'s; and $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

We can now say what J is: it's the quantum number associated with the magnitude of $\mathbf{J}$ (likewise for L and S). If the term "associated" seems too vague, here is the exact relationship: the magnitude of $\mathbf{J}$ is $\sqrt{J(J+1)}$.

We pick an arbitrary direction in space and call it the z-axis. M_J is the quantum number for the component of $\mathbf{J}$ along the z-axis. In other words, if $\mathbf{J} = (J_x, J_y, J_z)$ in some coordinate system, then M_J is the quantum number for J_z. Likewise for M_L and M_S. Note that $M_J = M_L + M_S$, but we have no such simple relation for J, L, and S; the magnitude of $\mathbf{J}$ depends of course on the angle between $\mathbf{L}$ and $\mathbf{S}$ as well as their magnitudes.

With our six angular momentum quantum numbers, we have enough ammunition to explain the Zeeman splitting (as we will see later). You may be wondering what happened to the quantum number n. Our six quantum numbers are not a complete set— that is, they are insufficient to fully specify the quantum state of the atom (or label the set of basis vectors, in modern terminology). But they are sufficient to explain the Zeeman effect.

Sommerfeld (and Bohr) knew nothing of spin, of course— that had to wait for Goudsmit and Uhlenbeck. Even so, Sommerfeld had concluded that he needed to augment his arsenal of quantum numbers, just from poring over spectra. Bohr and Sommerfeld supposed that in multi-electron atoms, the "inner core" of electrons interacted in some complex way with the outer electrons. That justified the introduction of more quantum numbers. (For historical accuracy, I note that Sommerfeld's set of quantum numbers differed from the "modern" set (J, S, etc.). The "modern" set is equivalent to Sommerfeld's, but conceptually clearer.)

Sommerfeld turned the problem of the anomalous Zeeman effect over to young Heisenberg, a student in his seminar. Heisenberg, struggling with it, introduced half-integer quantum numbers into the old quantum theory for the first time. Sommerfeld was shocked, and urged Heisenberg not to publish[10]. "If we know one thing about quantum numbers, we know they are integers!" Bohr, too, expressed displeasure. Pauli warned that today we allow half-integers, tomorrow it's quarter-integers, then eighths, sixteenths, and before you know it the quantum conditions have eroded away.

Selection Rules *How do the normal and anomalous Zeeman formulas ac-*

[9]ignoring any motion of the proton, a good approximation

[10]Alfred Landé independently came up with the same half-integer trick. He did publish, and so attached his name to the Landé g-factor.

count for the details of the Zeeman splitting?

I omitted many transition lines from figures 1 and 2 to avoid cluttering the diagram. Nature seems to share this taste for simplicity. Of the multitude of transition lines one might draw, most are forbidden by *selection rules*.

For figures 1 and 2, the appropriate selection rules assert that l changes by one unit, and m changes by at most one unit: $\Delta l = \pm 1$, $\Delta m = 0, \pm 1$. Conservation of angular momentum accounts for these rules, as hinted earlier. But before getting into that, let us explore the role of selection rules in the Zeeman effect.

As it happens, the normal and anomalous Zeeman effects call for different selection rules. Four selection rules apply to the normal Zeeman effect:

- $\Delta L = 0, \pm 1$.
- $\Delta M_L = 0, \pm 1$.
- $\Delta S = 0$.
- $\Delta M_S = 0$.

Plugging this into the formula $E^{\mathrm{mag}} = \text{constant}\, B(M_L + 2M_S)$, we get:

$$\Delta E^{\mathrm{mag}} = \text{constant}\, B \Delta M_L \tag{3}$$

Spin has dropped out completely! ΔM_L has only three possible values, thus the magnetic perturbation ΔE^{mag} splits each spectral line into three.

For example, consider transitions from $L = 2$ states to $L = 1$ states. Without the magnetic field, each $L = 2$ state belongs to a degenerate quintuplet of states (distinguished by the five possible values of M_L)— where "degenerate", you may recall, simply means that all states in the quintuplet have the same energy. Likewise, each $L = 1$ state belongs to a degenerate triplet. This degeneracy stems from the fact that the energy does not depend on M_L, when there is no field.

Had we no selection rules at all, we would have 15 possible transitions from the $L = 2$ quintuplet to the $L = 1$ triplet. The selection rule cuts this down to 9 (we can draw three lines to each state in the triplet). Since the quintuplet and the triplet are each degenerate, ΔE is the same for all 9 quintuplet-triplet transitions. We see a single spectral line from 9 transitions.

Now turn on the magnetic field. The degeneracy is lifted— the energy levels in the quintuplet separate slightly, as do those in the triplet (see figures 2). Group the 9 quintuplet-triplet transitions into three groups of three, according to the value of ΔM_L. Each group has the same ΔE^{mag}, and so 9 transitions give rise to three spectral lines.

For the anomalous Zeeman effect, different selection rules apply:

- $\Delta J = 0, \pm 1$.
- $\Delta M_J = 0, \pm 1$.

M_L and M_S fade from the picture. We will search for them again when we come to the Paschen-Back effect. (L and S are still with us, though.)

We want to plug this into formula 2 for the anomalous Zeeman E^{mag}. We don't need the full complexity of formula 2; the following is enough:

$$E^{\mathrm{mag}}_{JLSM_J} = f(J, L, S)M_J \tag{4}$$

The function $f(J, L, S)$ contains the complicated fraction we had before, plus constants (like the magnetic field strength B). Consider a transition from a (J, L, S, M_J) state to a (J', L', S', M'_J) state. Plugging in the selection rules, we get:

$$\Delta E^{\mathrm{mag}} = f(J, L, S)M_J - f(J', L', S')M'_J \tag{5}$$

As before, we first turn the field off, then on. Without the field, the energy of a state does not depend on M_J, and so each state belongs to a degenerate multiplet. We label the multiplet with the quantum numbers (J, L, S). Without the field, all transitions from one multiplet to another have the same ΔE—shining out one single spectral line.

Turn the field on, and this overloaded spectral line splits according to the different values of ΔE^{mag}. In contrast to normal Zeeman splitting, transitions with the same ΔM_J generally don't have the same ΔE^{mag}. Count the spectral lines, and you've counted the transitions (apart from accidental coincidences, which are rare).

For example, consider sodium, a historically important case. Sodium has two bright yellow spectral lines, known as D-lines (D_1 and D_2, close together).[11] In a weak magnetic field, the D_1 line splits in four, and the D_2 line splits in six.

"Scarcely credible!" wrote Lorentz of this discovery. Each D-line splits into an *even* number of lines. It would seem that some multiplet must contain an even number of states. But there are $2J + 1$ different M_J's for each value of J! Heisenberg and Landé unflinchingly ascribed half-integer values to J.

Here are the details. The D-lines come from three multiplets: a quartet and two doublets (see figure 3). The pair of quantum numbers (L, J) specify a multiplet, as can be seen from the figure— S is $\frac{1}{2}$ for all eight states. The D-lines are indicated by the slanted lines. The D_1-line actually encompasses all transitions from the higher to the lower doublet; the D_2-line encompasses all transitions from the quartet to the lower doublet.

Applying the selection rule for ΔM_J, we find that all four conceivable D_1 transitions are permitted; six out of the eight conceivable D_2 transitions are permitted.

[11] These lines are responsible for the color of sodium vapor lamps used on many highways.

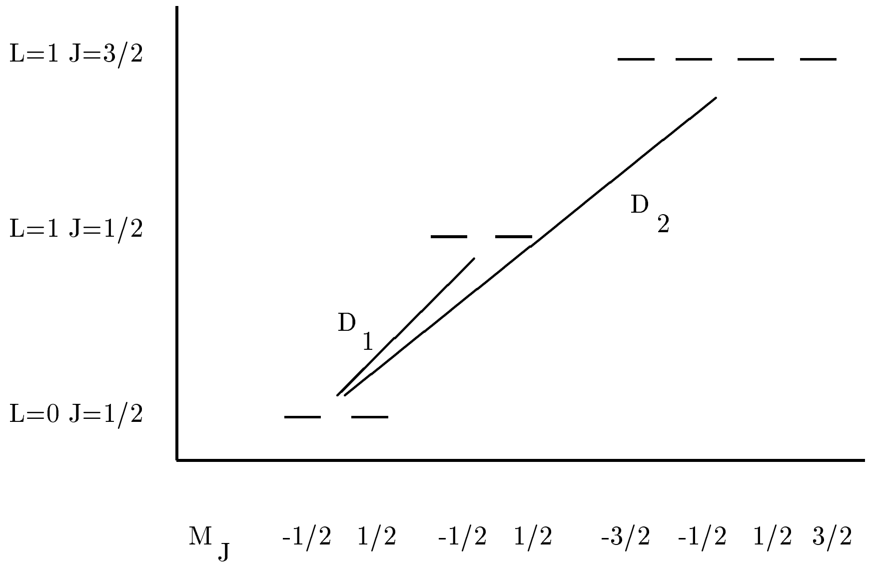

Figure 3: Sodium D-lines

Turn on the field, and these all separate. (See figure 4. To avoid cluttering the figure, the D_2-transitions are only partly drawn.) Hence the observed splitting.

I note finally that sodium is a "hydrogenic" atom. It has 11 electrons, one more than the noble gas neon. Ten electrons form a relatively inert "core"— they make up two complete "shells", which contribute collectively zero spin and zero orbital angular momentum (all electron spins are paired in the core.) The remaining electron is called a *valence* electron, and is far more loosely bound. A good approximation treats sodium as a system with a positively charged, immobile core, and a single orbiting electron— rather like hydrogen. The eight states we just discussed are (to a high degree of approximation) eight different states of the valence electron (which is why they all have $S = \frac{1}{2}$).

Where do the selection rules come from? Here is one way to view a transition:

$$\text{atom}_{\text{initial}} \rightarrow \text{atom}_{\text{final}} + \text{photon}$$

where "$\text{atom}_{\text{initial}}$" and "$\text{atom}_{\text{final}}$" refer of course to the initial and final states of the atom. $J_{\text{photon}} = 1$, so our rules for adding angular momenta say that

$$|J_{\text{final}} - 1| \leq J_{\text{initial}} \leq J_{\text{final}} + 1$$

in steps of 1, or in other words, $\Delta J = 0, \pm 1$ for the atom.

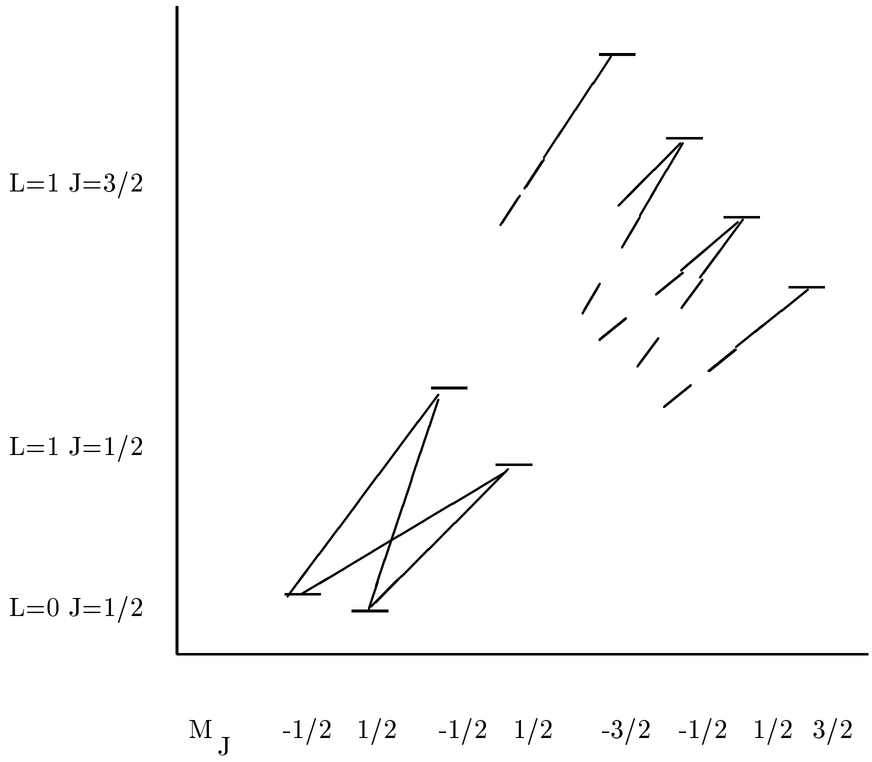

Figure 4: Zeeman Effect for Sodium

We must supplement the catalog of spin facts to get the rule for M_J. Our first "spin fact" told us that if we have a composite system:

$$C = A \oplus B$$

then $|J(A) - J(B)| \leq J(C) \leq J(A) + J(B)$, in steps of 1. We now add a new requirment:

$$|M_J(A) - M_J(B)| \leq M_J(C) \leq M_J(A) + M_J(B)$$

in steps of 1.

You may be wondering how to interpret these equations. What does it mean to say $C = A \oplus B$? Modern quantum mechanics gives a clear answer: tensor product. (So $\otimes$ would be a better notation than $\oplus$.) But a muddled view hews closer to history. The old quantum theory did offer derivations of the selection rules, via a strange brew of Maxwell's equations and other ingredients. On the other hand, the Paschen-Back effect presented a severe conceptual stumbling block. How does one set of selection rules fade away, and the other take over? Heisenberg discarded the derivations when it suited his purpose, a matter of some controversy at the time (see Cassidy's biography). So I leave a coherent discussion to the textbooks, and continue with cruder arguments.

To get the selection rules for ΔS and ΔM_S, we shift gears, and treat light as an electromagnetic wave, not photons. Also, we consider the reverse process, where an atomic state absorbs energy from incident light. An electron is so small (point-like, so far as we know today) that it feels at any moment a uniform electric field, and a uniform magnetic field. Classical physics says that magnetic effects are much smaller than electric effects, so we will ignore the former. (In modern parlance, we are "deriving" the selection rule for electric dipole transitions, and ignoring magnetic dipole transitions.)

The uniform electric field can shove the electron this way or that, but cannot flip it over. So $\mathbf{S}$ is unaffected, and hence $\Delta S = \Delta M_S = 0$. On the other hand, if the wavelength of light is short enough, the electric field will be shoving the electron one way in one part of its orbit, and another way for another part, and so may change its orbital angular momentum $\mathbf{L}$.

$\mathbf{J} = \mathbf{L} + \mathbf{S}$. We've just seen that $\mathbf{S}$ doesn't change, so $\mathbf{L}$ and $\mathbf{J}$ must change the same way. So the selection rules for L and M_L must be the same as those for J and M_J.

Magnetic Moments *Where do the normal and anomalous Zeeman formulas come from?*

First, a bit of classical physics. Imagine an "infinitesimally small" bar magnet—this is called a *magnetic dipole.* A vector (say $\boldsymbol{\mu}$) represents the strength and direction of the dipole; $\boldsymbol{\mu}$ is known as the *magnetic moment*.

Suppose we place the dipole in a magnetic field $\mathbf{B}$. The field will exert a torque on the dipole, trying to flip it into alignment (so $\mathbf{B}$ and $\boldsymbol{\mu}$ are parallel). Twisting the dipole out of alignment takes energy. In other words, the dipole has a potential energy that depends on its orientation with respect to the field. Classical physics says that this "magnetic" energy is given by $-\mathbf{B} \cdot \mu$.

A tiny current loop acts like a dipole, with a magnetic moment perpendicular to the plane of the loop. The current loop possesses angular momentum (after all, the current consists of moving charges, which have mass). According to classical physics, the magnetic moment is proportional to the angular momentum: $\mu =$ constant times $\mathbf{J}$. The constant depends on the charge/mass ratio of the moving charges making up the current.

We now apply all this to an electron in an atom in a magnetic field. Thanks to its orbital angular momentum $\mathbf{l}$, it has a "magnetic" energy proportional to $\mathbf{B} \cdot \mathbf{l}$; likewise, the spin contributes energy proportional to $\mathbf{B} \cdot \mathbf{s}$. Adding up the $\mathbf{l}$'s and $\mathbf{s}$'s for all the electrons in the atom, we might expect a formula like this for the total "magnetic" energy:

$$E^{\text{mag}} = \text{constant}\,\mathbf{B}\cdot(\mathbf{L}+\mathbf{S}) = \text{constant}\,\mathbf{B}\cdot\mathbf{J}$$

Instead, the correct formula is:

$$E^{\text{mag}} = \text{constant}\,\mathbf{B}\cdot(\mathbf{L}+2\mathbf{S}) \tag{6}$$

Let us accept this formula, noting but not discussing the mysterious factor 2. If we pick the z-axis in the direction of $\mathbf{B}$, then the dot product simply pulls out the z-component of $\mathbf{L}$ and $\mathbf{S}$, times the magnitude of $\mathbf{B}$. The quantum numbers M_L and M_S give the values of these components. So we get the "normal Zeeman" formula:

$$E^{\text{mag}}_{LM_LSM_S} = \text{constant}\,B(M_L+2M_S)$$

or equivalently:

$$E^{\text{mag}}_{LM_LSM_S} = \text{constant}\,B(M_J+M_S) \tag{7}$$

since $M_J = M_L + M_S$.

Without the enigmatic 2, we'd simply have M_J on the right hand side of formula 7. Without M_S on the right hand side, the normal and anomalous Zeeman formulas would be the same— as we will see.

How does the electron maintain its orientation in the magnetic field— why doesn't it snap into alignment? Or in "old quantum" terms, how can we have stationary states with $\mathbf{S}$ (or $\mathbf{L}$) out of alignment with the field? This was a legitimate question for the old quantum theory, and it had a good answer. The spinning electron acts like a gyroscope. The field exerts a torque on the electron, and like any gyroscope, the electron precesses. (The same argument holds for $\mathbf{L}$.)

We can get the "anomalous Zeeman" formula from formula 7 with a little vector algebra, and a crucial assumption:

> Assume that in any stationary quantum state, $\mathbf{S}$ and $\mathbf{L}$ each precess around $\mathbf{J}$, so that the projection of $\mathbf{S}$ on $\mathbf{J}$ doesn't change, but the component of $\mathbf{S}$ perpendicular to $\mathbf{J}$ rotates uniformly (ditto for $\mathbf{L}$).

This is called *spin-orbit coupling*. Only without an external magnetic field is it strictly true. It is approximately true if the external field is weak enough. It breaks down in a strong magnetic field, a phenomenon known as spin-orbit decoupling.

Where does spin-orbit coupling come from? Naively expressed, from a torque trying to make $\mathbf{L}$ and $\mathbf{S}$ point opposite to each other (i.e., anti-parallel). Goudsmit and Uhlenbeck imagined it this way. Suppose there is only a single electron. If we ride along an electron's orbit right next to it, but not spinning ourselves, we will appear to see the nucleus orbiting *us* (shades of Ptolemy!) Effectively we have a positive current circulating around us; this will generate a magnetic field proportional to $\mathbf{L}$. The electron has a magnetic moment proportional to $\mathbf{S}$, and hence we get an energy term proportional to $\mathbf{L} \cdot \mathbf{S}$. (Because of the positive charge on the nucleus, the signs come out so that spin-orbit coupling trys to make $\mathbf{L}$ and $\mathbf{S}$ anti-parallel.)

Life is more complicated in multi-electron atoms. $\mathbf{L} \cdot \mathbf{S}$ is only an approximation, applicable when the electron-electron forces dominate the spin-orbit forces. In this case, one can approximate the totality of electrons as a single electron "smear", with total orbital angular momentum $\mathbf{L}$, and total spin $\mathbf{S}$. For "light" atoms, this works pretty well. For "heavy" atoms, a different approximation works better.

A classical model has emerged. The vectors $\mathbf{L}$ and $\mathbf{S}$ feel a torque trying to make them anti-parallel; when the magnetic field $\mathbf{B}$ is turned on, $\mathbf{L}$ and $\mathbf{S}$ each feel a torque trying to align them with $\mathbf{B}$. We have, in other words, "perturbation terms" occuring somewhere in the total energy formula:

$$c_1 \mathbf{L} \cdot \mathbf{S} + c_2 \mathbf{B} \cdot (\mathbf{J} + \mathbf{S}) \tag{8}$$

where c_1 and c_2 are constants, and $\mathbf{S}$ occurs on the right hand side only because of the mysterious factor 2 in formula 6.

Classical mechanics would next try to compute the "trajectories" of the vectors $\mathbf{L}$ and $\mathbf{S}$ (i.e., how they vary with time). If $\mathbf{B} = 0$, then $\mathbf{J}$ is constant (conservation of angular momentum), and $\mathbf{L}$ and $\mathbf{S}$ precess about $\mathbf{J}$. If the field is weak, then (to a good approximation) $\mathbf{J}$ precesses about $\mathbf{B}$, while $\mathbf{L}$ and $\mathbf{S}$ still precess about $\mathbf{J}$. $\mathbf{J}$ is no longer constant, since it is subject to an external torque. If the field is strong, then $\mathbf{L}$ and $\mathbf{S}$ precess about $\mathbf{B}$ (again to a good approximation), because the magnetic torque swamps the spin-orbit torque.

The "old quantum" prescription says to find the energy levels of these stationary states.

As it happens, the energy of a stationary state depends primarily on terms other than those in formula 8. We have terms analogous to the R/n^2 of hydrogen; then terms depending on the magnitude of $\mathbf{L}$ and $\mathbf{S}$ (i.e., terms depending on L and S); only after that does formula 8 come into play.

The spin-orbit coupling contributes a perturbation that depends on the angle between $\mathbf{L}$ and $\mathbf{S}$ (and hence indirectly on the magnitude of $\mathbf{J}$). So when $\mathbf{B} = 0$, we expect to find bunches of energy levels with the same L and S grouped closely together, separated slightly by an energy difference depending on J. This is an aspect of the famous *fine structure* of the spectrum.[12] If $S = 0$, then the fine structure disappears. (Low-lying energy levels of atoms with even numbers of

[12] The sodium D-lines provide an example. The quartet and the higher doublet both have $L = 1$ and $S = \frac{1}{2}$; their J's are $\frac{3}{2}$ and $\frac{1}{2}$.

electrons usually have $S = 0$.)

Without a field, we have complete rotational symmetry in the energy formula. The spin-orbit perturbation depends only on the *magnitude* of $\mathbf{J}$, not on its direction (on J, not on M_J). With a field, the symmetry is broken and the degeneracy lifted.

We have perturbations depending on $\mathbf{B} \cdot \mathbf{L}$ and $\mathbf{B} \cdot \mathbf{S}$. With a strong field (or if $S = 0$), we confidently replace these dot products with BM_L and BM_S, for the dot products don't change and must be quantized (i.e., they are stationary). We have seen already how this leads to the normal Zeeman effect. Note also another regularity: spectra without fine-structure have $S = 0$, and hence display normal Zeeman splitting.

With a weak field, we express the perturbation in terms of $\mathbf{B} \cdot \mathbf{J}$ and $\mathbf{B} \cdot \mathbf{S}$. The first term we replace with BM_J. Were this all, we would again see a normal Zeeman effect, because ΔM_J and ΔM_L satisfy the same selection rule.

Turn to the $\mathbf{B} \cdot \mathbf{S}$ perturbation. Because $\mathbf{S}$ precesses about $\mathbf{J}$, the *average* value of $\mathbf{B} \cdot \mathbf{S}$ is the same as

$$\mathbf{B} \cdot \mathbf{S}_{\|}$$

where $\mathbf{S}_{\|}$ is the projection of $\mathbf{S}$ on $\mathbf{J}$. The dot product $\mathbf{B} \cdot \mathbf{S}_{\|}$ doesn't change, so it must be quantized. $\mathbf{S}_{\|}$ is some fraction of $\mathbf{J}$, so this dot product must be some fraction of M_J. What fraction? If we knew the lengths of $\mathbf{L}$ and $\mathbf{S}$ and the angle between them, we'd clearly have enough information to specify the geometry completely. Instead of the angle, we could also use the length of $\mathbf{J}$. So we expect the fraction to depend on L, S, and J. And that, as we've seen, accounts for the anomalous Zeeman effect.

The rest is vector algebra. You may skip to the next section if you like— no new concepts appear.

First derive a formula for $\mathbf{L} \cdot \mathbf{S}$, in terms of quantum numbers. We have:

$$\mathbf{J} \cdot \mathbf{J} = (\mathbf{L} + \mathbf{S}) \cdot (\mathbf{L} + \mathbf{S}) = \mathbf{L} \cdot \mathbf{L} + \mathbf{S} \cdot \mathbf{S} + 2\mathbf{L} \cdot \mathbf{S}$$

Rearrange this equation and replace $\mathbf{J} \cdot \mathbf{J}$ with $J(J+1)$ (and likewise for $\mathbf{L}$ and $\mathbf{S}$), to get:

$$\mathbf{L} \cdot \mathbf{S} \to \frac{J(J+1) - L(L+1) - S(S+1)}{2} \tag{9}$$

where I've used an arrow instead of an equal sign to indicate "quantization"— classical quantities on the left, quantum numbers on the right. From the modern perspective, the left hand side is an operator, and the right hand side is one of its eigenvalues. (Formula 9 looks tantalizingly close to the Landé g-factor, but we're not quite there yet.)

Next derive a formula for $\mathbf{S}_{\|}$, the projection of $\mathbf{S}$ on $\mathbf{J}$:

$$\mathbf{S}_{\|} \quad = \quad \frac{\mathbf{J} \cdot \mathbf{S}}{\mathbf{J} \cdot \mathbf{J}} \mathbf{J}$$

$$= \left(\frac{\mathbf{L}\cdot\mathbf{S}}{\mathbf{J}\cdot\mathbf{J}}+\frac{\mathbf{S}\cdot\mathbf{S}}{\mathbf{J}\cdot\mathbf{J}}\right)\mathbf{J}$$

"Quantizing" the term in parentheses:

$$\frac{J(J+1)-L(L+1)-S(S+1)}{2J(J+1)}+\frac{S(S+1)}{J(J+1)}=\frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)}$$

In other words, $\mathbf{S}_\parallel$ is this fraction times $\mathbf{J}$. Taking the dot product with $\mathbf{B}$, we got previously:

$$E^{\mathrm{mag}} = \text{constant}\, B(M_J + M_S)$$

Replace the term BM_S with the "average" value of $\mathbf{B}\cdot\mathbf{S}$, which is $\mathbf{B}\cdot\mathbf{S}_\parallel$, which is the above fraction times BM_J:

$$E^{\mathrm{mag}} = \text{constant}\, B(M_J + \frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)}M_J)$$

And this reduces to our anomalous Zeeman formula, formula 2.

The Paschen-Back Limit *How does the anomalous Zeeman formula turn into the normal Zeeman formula in the Paschen-Back limit of strong magnetic fields?*

The previous section puts classical physics in a rather positive light. A thoroughly classical picture appears to account for fine structure, normal Zeeman splitting, anomalous Zeeman splitting, and the passage from one to the other as the field strength increases.

On closer examination, we find cracks in the picture. Several non-classical ingredients played a crucial role:

1. The Bohr-Sommerfeld "quantization" prescription.
2. The selection rules.
3. The spin one-half of the electron.
4. The "mysterious factor 2".

The old quantum theory had problems with all of these:

1. The quantization rules apply only to "stationary" states, but just what constitutes a stationary state? A stationary state is not absolutely stable, since it can make transitions.
2. I've noted already the muddle over selection rules as we pass from the anomalous to the normal Zeeman effect.

3. Heisenberg and Landé had half-integer spin forced upon them by counting spectral lines. Now the quantum number L assumes only integer values; Sommerfeld had a derivation of sorts for this fact (which I have not given). Why should this derivation work for L but not for S? Neither Heisenberg nor anyone else had an answer at the time.

4. This is related to the previous point, but is a distinct issue. The ratio of magnetic moment to angular momentum depends on what kind of momentum we're dealing with: orbital or spin.

The question of "stationary" states comes to the fore when we look at the Paschen-Back effect. Modern progress in chaos theory has clarified the subject further. For weak fields, we have one kind of approximately stationary state ($\mathbf{L}$ and $\mathbf{S}$ precess about $\mathbf{J}$); for strong fields, a different kind ($\mathbf{L}$ and $\mathbf{S}$ precess about $\mathbf{B}$.) For intermediate strength fields, the motion becomes chaotic (classically speaking). The Bohr-Sommerfeld quantization procedure throws up its hands when faced with chaotic motion. Yet quantization does not cease when the motion becomes chaotic, as we can tell by gazing at spectral lines.

Quantum mechanics broke the logjam. The stimulus for the fundamental shift in viewpoint did not come from the Zeeman effect, it is true. But the new approach unlocked this and many other mysteries within a year or two.

The next few sections sketch the resolution of these problems. I will do a pretty good job for (1) and (2), an overbrief treatment of (3), but I will barely touch on (4), regretfully leaving the real explanation to the standard textbooks.

Quantization Schrödinger titled his first great paper on quantum mechanics "Quantization as an Eigenvalue Problem". I have alluded to the meaning of this several times already. A state specified by v has quantum number q for operator Q if v satisfies the eigenvalue equation $Qv = qv$. Borrowing the imagery of the old quantum theory, we say that Q *has the value* q in the state specified by v.

Other jargon says that Q is "sharp" or "definite" in the state v. According to the quantum theory of measurement, if we prepare an ensemble of systems, all in the state v, and measure Q in each one, we will always get the value q. But if v is not an eigenvector of Q, then we will various eigenvalues of Q with different probabilities. There is statistical spread, and the value of Q for v is not "sharp".

What about "stationary"? Here we must bring in another concept I've mentioned before: the energy operator governs the time evolution of a quantum system. A state v is absolutely stationary (does not change at all) if and only if it is an eigenstate of the energy operator E. To be a touch more precise: suppose $Ev = \omega v$. Then v evolves like so:

$$v(t) = e^{i\omega t}v(0)$$

Remember that v and cv specify the same state for any non-zero complex c. So the state doesn't change. (I am also assuming that the energy operator does not change with time. A changing E calls for a different equation for $v(t)$.)

Bohr's states are "almost" stationary. Perturbation theory deals with this ambiguity. Express the energy operator as a sum of two parts:

$$E = H + V$$

where V, the perturbation, is "small" relative to H. Eigenstates of H will not usually be eigenstates of E, but they will change "slowly" because V is "small"[13]. Start the system in an eigenstate of H, $v(0)$ say. Wait awhile. The system is now in a superposition of eigenstates of H— say, for simplicity, $v(1) = c_1 v(0) + c_2 u$. Here u is another eigenvector of H, and c_1 and c_2 are complex numbers. If we measure H, we have a certain probability of finding the system in the state u.

In other words, the system has made a transition from one state to another, under the influence of a perturbation.

Bohr's transition picture is a special case of this. The perturbation V is the electromagnetic term, representing the interaction between the electromagnetic field and the atom. In other words, V is due to the ability of the atom to make transitions by absorbing or emitting a photon. Everything else is stuffed into H: the attraction of the nucleus, spin-orbit coupling, the constant magnetic field of the Zeeman effect. (Classical electromagnetism gives an unambiguous way to separate this constant magnetic field from the travelling field of light.) Bohr's stationary states are the eigenstates of H.

What has become of Bohr's notion of quantum number? For Bohr and Sommerfeld, stationary states had quantum numbers. Translated into Hilbert space language, this asserts that eigenvectors of H are eigenvectors of all other operators of physical importance. Alas, life is not so easy. If operators Q and H commute, one can generally pick a basis of eigenvectors of both Q and H. (Trivial exercise: prove the converse.) Some phenomena do submit to such an approach. The Paschen-Back effect does not— as we will soon see.

Selection Rules Redux We start in a stationary state $v(0)$. We wait a bit. Under the influence of the electromagnetic perturbation V, v evolves to a state which is a superposition of $v(0)$ and other stationary states u_j. What can we say about the possible u_j?

At first you might guess that one could not say anything without a detailed study of V. Say $\{u_j\}$ is a basis of eigenvectors of H. Expand $v(t)$ out in this

[13] In particular cases, "small" and "slowly" can be given more definite meaning. For example, we may be able to write E as a power series in something; V might contain the second and higher order terms.

basis:

$$v(t) = \sum_j c_j(t) u_j$$

It turns out that for small positive t, $c_j(t) \approx 0$ unless $Vv(0)$ contains a non-zero u_j component. Say we set $v(0) = u_i$, and expand out Vu_i in the $\{u_j\}$ basis—$Vu_i = \sum_j a_{ij} u_j$. We must have $a_{ij} \neq 0$ to have a significant chance of the transition $u_i \to u_j$.

Let's rephrase this without so many indices. V has a matrix representation in the basis of eigenvectors of H. "Immediate" transitions between "stationary states" (i.e., eigenvectors of H) come from non-zero off-diagonal entries in the V matrix.

Over longer periods of time, we can have transitions through intermediate states: $u_i \to u_j \to u_k$. If you work through the math, you will find out that you are computing V^2, V^3, etc., for these indirect transitions; the power series for $\exp(iV)$ makes an appearance in the final result.

The moral of this tale is that *zero entries in the V matrix correspond to forbidden transitions.* A selection rule translates into an assertion about the form of the V matrix. And so, it appears, we need a detailed study of V to determine the selection rules.

Our crude "derivation" of the selection rules for S indeed depended on the "mechanism" of light. But for J, we appealed to a general physical principle, conservation of angular momentum. Can one not translate this argument into quantum mechanics?

One can. The thread runs thus: V must be invariant under all spatial rotations, for electromagnetism does not single out any preferred direction in space. So the group of spatial rotations, $SO(3)$, must play a special role. At this point group representation theory takes over, and out pops the selection rules for J and M_J. (You do need some additional assumptions I won't spell out.)

Let's take a last look at the Paschen-Back effect, using all we've learned. Where do the J, M_J selection rules leave off and the L, M_L and S, M_S selection rules take over, as we increase the magnetic field?

Both sets of selection rules hold throughout! The trick is picking the right basis. To understand this, we must consider again the combined influence of the spin-orbit and magnetic perturbations.

The operator H looks like this:

$$H = H^0 + c_1 \mathbf{L} \cdot \mathbf{S} + c_2 \mathbf{B} \cdot (\mathbf{L} + 2\mathbf{S})$$

or equally well:

$$H = H^0 + c_1 \mathbf{L} \cdot \mathbf{S} + c_2 \mathbf{B} \cdot (\mathbf{J} + \mathbf{S})$$

H^0 contains terms representing the attraction of the nucleus, and perhaps other refinements for complex atoms. The next two terms are the spin-orbit coupling and the effect of the magnetic field.

$\mathbf{B}$ is just a conventional 3-space vector, but $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{J}$ are all "operator vectors". That is, for any coordinate system, we have $\mathbf{L} = (L_x, L_y, L_z)$, where L_x, L_y, and L_z are all operators. Ditto for $\mathbf{S}$ and $\mathbf{J}$.

Pick the z-axis in the direction of the magnetic field. Several operators now demand a role:

$$
\begin{array}{rcl}
\mathbf{J}\cdot\mathbf{J} & = & J_xJ_x + J_yJ_y + J_zJ_z \\
\mathbf{L}\cdot\mathbf{L} & = & L_xL_x + L_yL_y + L_zL_z \\
\mathbf{S}\cdot\mathbf{S} & = & S_xS_x + S_yS_y + S_zS_z \\
\mathbf{L}\cdot\mathbf{S} & = & L_xS_x + L_yS_y + L_zS_z \\
J_z, & L_z, & S_z
\end{array}
$$

some with eigenvalues we've come to know and love:

$$
\begin{array}{rcl}
\mathbf{J}\cdot\mathbf{J} & \to & J(J+1) \\
\mathbf{L}\cdot\mathbf{L} & \to & L(L+1) \\
\mathbf{S}\cdot\mathbf{S} & \to & S(S+1) \\
J_z & \to & M_J \\
L_z & \to & M_L \\
S_z & \to & M_S
\end{array}
$$

Alas, these operators do not all commute. It turns out that $\mathbf{L}\cdot\mathbf{L}$ and $\mathbf{S}\cdot\mathbf{S}$ commute with each other and all the rest, and hence with H. For this reason, L and S are "good" quantum numbers: we can pick stationary states (in the Bohr sense) that have sharp values for L and S.

Finding additional "good" quantum numbers proves more frustrating. $\mathbf{J}\cdot\mathbf{J}$ commutes with J_z and with $\mathbf{L}\cdot\mathbf{S}$, but not with L_z or S_z (as it happens). Drop the $\mathbf{B}\cdot\mathbf{S}$ term from H, and we have J and M_J as good quantum numbers. Alternately, drop the spin-orbit coupling term, and, as luck will have it, M_L and M_S become good quantum numbers.

So we have a choice of bases. With no field, we can find stationary states with sharp values of (L, S, J, M_J). With no spin-orbit coupling, we can demand sharp values for (L, S, M_L, M_S).

Start with the (L, S, J, M_J) basis, and turn on a weak field. The states are now only approximately stationary (even in Bohr's sense). The stronger the field, the less accurate the approximation. But the matrix elements for V, in this basis, strictly obey the J and M_J selection rules, no matter how strong the field.

For a very strong field, the (L, S, M_L, M_S) basis consists of approximately stationary states (i.e., near-eigenvectors of H). The M_L and M_S selection rules hold strictly with respect to this basis, no matter how weak the field. (The L and S selection rules hold for either basis.)

Just to hose away the last traces of the muddle: how can we reconcile $\Delta M_S = 0$ with the formulas $M_S = g(J, L, S)M_J$ and $\Delta M_J = 0, \pm 1$? (We needed the latter two formulas for the anomalous Zeeman effect.) Answer: Suppose an atom makes a transition from $|M_J = \uparrow\rangle$ to $|M_J = \downarrow\rangle$ (using the Dirac $|\rangle$ notation for state-vectors, plus the "colloquial" abbreviations $\uparrow = \frac{1}{2}$, $\downarrow = -\frac{1}{2}$). The atom begins and ends in an eigenstate of J_z. Each eigenstate is a "blend" of eigenstates of S_z, say:

$$\begin{aligned} |M_J = \uparrow\rangle &= a\,|M_S = \uparrow\rangle + b\,|M_S = \downarrow\rangle \\ |M_J = \downarrow\rangle &= c\,|M_S = \uparrow\rangle + d\,|M_S = \downarrow\rangle \end{aligned}$$

The selection rule for M_S holds, in the sense that $|M_S = \uparrow\rangle$ makes a transition only to itself; likewise for $|M_S = \downarrow\rangle$. But the coefficients a, b, c, and d change. The M_S in the equation $M_S = g(J, L, S)M_J$ is an "average" M_S, and depends on these coefficients. (More precisely, it is the expectation value of S_z for a state which is an eigenstate of J_z, not of S_z.) The M_S of the selection rule is the quantum number of an eigenstate of S_z.

We have already seen this resolution foreshadowed in the classical treatment, when we obtained the "average" z-component of $\mathbf{S}$ by taking the z-component of $\mathbf{S}_\parallel$. But classical mechanics lacks the notion of "blended" states, and so is ill-equipped to pass smoothly from the weak field to the strong field regime.

So much for the Zeeman effect. Let us punctuate the tale with an anecdote. A friend ran into Heisenberg on the streets of Cophenhagen, around 1920; Heisenberg had a grim expression. "Cheer up, Werner, things can't be that bad!" Replied Heisenberg, "How can one be cheerful when one is thinking about the anomalous Zeeman effect?"

Spin One-half, and the Mysterious Factor 2 Lie group and Lie algebra representation theory fit flawlessly into the structure of quantum mechanics, and the most satisfying explanation of the mysteries of spin one-half lie in this connection. Careful exposition of this material belongs to standard textbooks. But I must at least justify my remark that half-integer spin is fundamentally non-classical.

Let us start with a quantum mechanical system which has n states. The Hilbert space for this system is then isomorphic to $\mathbf{C}^n$. Suppose the system "inhabits" ordinary physical Euclidean 3-space, that is, we can picture the physical processes of the system as taking place in 3-space. Vague though this statement is, something precise will come out of it. Any classical physical system (e.g., a spinning ball) "inhabits" 3-space in this (as yet) fuzzy sense.

Rotate the system in 3-space. Or if you prefer, rotate the coordinate axes used to describe the system. Either way, we have a transformation from one state of the system to another. (Physicists like to distinguish "active" from "passive" transformations, along the lines of preference just mentioned, but we won't need to be so exact.) We will assume this transformation can be represented by a linear operator on the Hilbert space of the system.

In fact, no generality is lost by assuming the linear operator is unitary— one can prove this. Even more: since our Hilbert space is isomorphic to $\mathbf{C}^n$, the operator has a matrix representation, and it has been proved that one can arrange for the determinant of the matrix to be 1. In short, we have a unitary unimodular representation of the Lie group $SO(3)$ on $\mathbf{C}^n$:

$$\sigma : SO(3) \to SU(n)$$

This is the precise statement that emerges from our vague notion of "inhabiting 3-space".

Group representation theory now grinds away. It tells us first, that σ is a direct sum of irreducible representations, and next, that an irreducible representation exists if and only if the dimension of the target space is odd (in which case the representation is essentially unique).

So we now have:

$$SO(3) \to SU(H_1) \oplus \ldots \oplus SU(H_r)$$

where $\Sigma_i H_i$ is a direct sum decomposition of the Hilbert space of our system, and $SU(H_i)$ is the group of unitary unimodular transformations on H_i; also, each H_i has odd dimension.

Any state in subspace H_i can be rotated into any other state in subspace H_i. But rotation never mixes different subspaces together: a state in subspace H_i remains in that state. So it seems natural to concentrate on the irreducible representations.

Associated with the Lie group $SO(3)$ is the Lie algebra $so(3)$, generated by three Lie algebra elements; we can picture these generators as "infinitesimal" rotations about the x, y, and z axes (as did Lie), or as angular velocities about the coordinate axes.

Suppose $\sigma : SO(3) \to SU(H)$ is an irreducible representation, and suppose H has dimension $2j+1$. Then σ induces a map from $so(3)$ to the Lie algebra $su(H)$, which just happens to consist of all the anti-Hermitian traceless operators on

H. The three generators of $so(3)$ map to three operators I will label iJ_x, iJ_y, and iJ_z— that factor of i makes J_x, J_y, and J_z Hermitian.

The generators of $so(3)$ "look like" angular velocities, so J_x, J_y, and J_z are good candiates for the angular momentum operators. Ultimately this comes down to a matter of definition. Let us make this identification without further ado.

Finally, the operator $J_xJ_x + J_yJ_y + J_zJ_z$ commutes with all of $SU(H)$ (as can be shown by direct calculation, or more cleverly), and so by Schur's lemma, is a constant times the identity matrix. And now the punchline: it turns out that this operator is $j(j+1)I$, where j (you may recall) is defined by the relation $\dim H = 2j+1$.

In one sense we should be well satisfied. Our original direct sum decomposition of the n-dimensional Hilbert space of the system simply decomposed that space into states with a definite magnitude for the angular momentum— that is, with definite quantum number j. On the other hand, group representation theory has told us that j must be an integer and $\dim H$ is odd. And we want $j = \frac{1}{2}$ for an electron.

But now something curious appears. $SO(3)$ has a double-cover, the Lie group $SU(2)$. Their Lie algebras are of course the same. Perhaps we pick up some additional representations by starting with $SU(2)$?

Indeed we do. There is a unique irreducible representation $SU(2) \to SU(k)$ for every integer k; for reasons that should be obvious by now, we set $2j+1 = k$, and use j in place of k in all the formulas. The $SU(2)$ representation factors through the $SO(3)$ representation precisely when j is an integer: $SU(2) \to SO(3) \to SU(2j+1)$.

This is as far as I will pursue the mathematics. Our Hilbert spaces are still rather bloodless. Schrödinger conjured up lovely images of wavefunctions spreading through 3-space, which is to say he picked a representation for the Hilbert space which "inhabits 3-space" in our sense. Naturally the Schrödinger wavefunctions will not serve to represent a spin one-half particle like an electron. Pauli solved this puzzle, scant months after the invention of quantum mechanics. Relativity led to further paradoxes. To Dirac belongs the glory of resolving these. But thereby hangs another tangled tale of history...

I will ramble a bit further though on classical imagery versus quantum reality, just with regard to this question of spin. The groups $SU(2)$ and $SO(3)$ have the same Lie algebra, and as we've seen, the Lie algebra begets the angular momentum operators. This above all is why classical pictures can carry us so far, even for half-integer spin. Push it far enough though, and any classical picture will finally break down, for $SU(2)$ and $SO(3)$ are different groups.

In some sense, "the mysterious factor 2" stems from this difference. Because $SU(2)$ is the double cover of $SO(3)$, factors of 2 appear in the formula for the covering projection. These factors ultimately find their way into the formula for the ratio of the electron's magnetic moment to its spin angular momentum. The orbital motion of an electron however "inhabits 3-space", and so no factors of 2 appear.

Helium Having dwelt lovingly upon the Zeeman effect, I will pass more fleetingly over some of the other applications of spin. I wish only to give an inkling of the pervasiveness of the concept.

Close study of the spectrum of helium revealed a striking feature. There are two systems of energy levels, which do not interact with each other— transitions between the systems do not exist. The bottom level of the first system is well below the bottom level of the second system. I omit a detailed catalog of other regularities.

Born, Heisenberg, Pauli, and other applied more and more recondite techniques from classical mechanics to the problem of the helium spectrum (specifically celestial mechanics, including Poincaré's work). All in vain. As I've already noted, helium poses one of the simplest three-body problems in quantum mechanics, and the old quantum theory just couldn't handle it.

Heisenberg solved the problem after the invention of quantum mechanics. The two electrons each have spin-$\frac{1}{2}$, so the total spin S is either 0 or 1. States with $S = 0$ are called *singlet states*, states with $S = 1$ are *triplet* states (recall again the formula $2S+1$ for the number of possible values for M_S). The selection rule $\Delta S = 0$ accounts for the division of energy levels into two subsystems. If $S = 1$, then the electrons have the same spin, and cannot both occupy the bottom level in a term scheme like figure 1. If $S = 0$, they can. Thus the bottom singlet state is significantly lower than the bottom triplet state. The bottom triplet state is sometimes called *metastable*, for processes other than photon absorption or emission can (slowly) change a triplet state into a singlet state. (The $\Delta S = 0$ selection rule applies strictly only to electric dipole radiation transitions.)

Heisenberg noted that the hydrogen molecule H_2 should also have a similar "double spectrum", corresponding to $S = 0$ and $S = 1$ states. (Here we ignore the spins of the protons and count only the spins of the two electrons.) In a sense, H_2 has two "allotropic forms", dubbed parahydrogen and orthohydrogen. These two forms were subsequently discovered by experimentalists. The Nobel committee cited this work when awarding Heisenberg the Nobel prize.

Phosphorescence First, let's distinguish phosphorescence from fluorescence. A fluorescent paint glows under a UV lamp, but stops glowing as soon as the lamp is turned off. A phosphorescent paint keeps glowing for a while.

Phosphorescent substances have the ability to store up light and release it gradually. The notion of a metastable state explains this. If the molecules of the substance can get from the ground state to a metastable state, and if the metastable

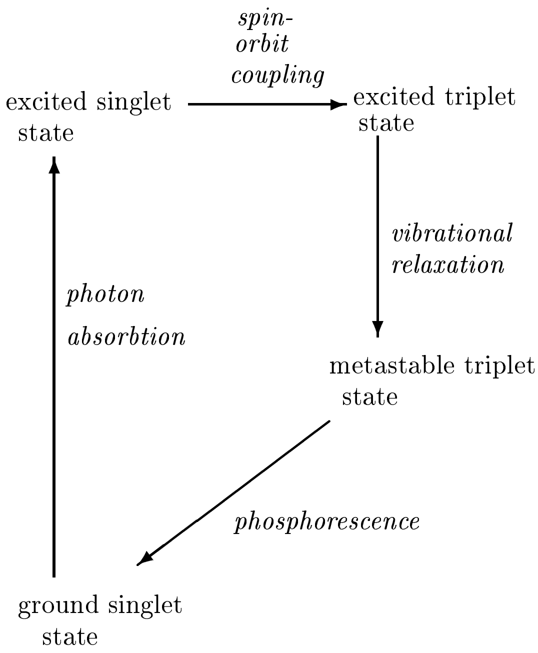

Figure 5: Phosphorescence

state can slowly decay back to the ground state via photon emission, then we have phosphorescence.

Typically, the metastable state is a triplet state, and the ground state is a singlet state. Ground state molecules absorb photons and go to excited singlet states (see figure 5). Most of them immediately hop right back to the ground state, emitting a photon, but non-radiative processes take a few to a less energetic triplet state. Once these molecules get to the lowest triplet state, they are stuck there, at least for a while. Some low probability process accomplishes the triplet-singlet conversion, and the molecules slowly leak out light.

What are the singlet-triplet and triplet-singlet processes? I mention one possibility of several. If the molecule contains two unpaired electrons, and each is subject to spin-orbit coupling— but with different strengths— then $\mathbf{s}_1$ and $\mathbf{s}_2$ will precess at different rates, and the magnitude of the combined vector $\mathbf{s}_1 + \mathbf{s}_2$ will flip-flop between $S = 0$ and $S = 1$.

Complicating matters a little, the energy of a molecule consists of many parts. For example, the vibration and rotation of the "nuclear framework" stores energy. Suppose the excited singlet state is close in energy to a triplet state. After the molecule makes the singlet-triplet transition, it may shed energy into the environment through the vibrational modes (*vibrational relaxation*). Finally it arrives at the bottom triplet state. Eventually it decays back to the ground state via a triplet-singlet transition.

Both the singlet-triplet and triplet-singlet transitions violate the $\Delta S = 0$ selection rule. This rule applies absolutely only for pure electric dipole transitions, so the glow-in-the-dark T-shirts don't violate the laws of universe. Still, transitions that violate it are apt to be slow.

One last fact completes the explanation. A result of perturbation theory says that the probability of a transition between state u_i and state u_j is determined by two factors: the matrix element connecting the two states (the "strength" of the coupling), and the difference in energy levels. It is, in general, easier to make a transition when the energies are close together. Thus the singlet-triplet transition goes faster than the triplet-singlet transition.

Loose Ends Why $j(j+1)$, instead of j^2? Well, group representation theory tells us so!— as already mentioned in the section titled "Spin One-half". If you consult the computations in a textbook, you will find it all hinges on the commutators of the Lie algebra. You may also find it amusing to compute the "average value of m^2" (m being the projection on the z-axis, as usual) like so:

$$\frac{(-j)^2 + (-j+1)^2 + \ldots + (j-1)^2 + j^2}{2j+1}$$

and verify that $j(j+1)/3$ results. Now argue that for any vector $\mathbf{v}$, $|\mathbf{v}|^2 = v_x^2 + v_y^2 + v_z^2$, so the average value of m^2 should be one-third the magnitude of the angular momentum vector.

How do we reconcile the verdict of group representation theory with the classical value j^2? I suggested, loosely, that "$SO(3)$ is classical, $SU(2)$ is quantum". The eigenvalue $j(j+1)$ applies to representations of $SO(3)$ as well as $SU(2)$.

The culprit here is the finite dimensionality of the representation. To call j the "classical value" for the angular momentum is misleading. Let us return to "unnatural units", where $\hbar$ is small. The "classical" magnitude for angular momentum is $j\hbar$, with large j. We get the "classical limit" by simultaneously letting $j \to \infty$ and $\hbar \to 0$, while keeping the product constant. In the classical limit, the quantum expression $\sqrt{j(j+1)}\hbar$ simplifies to $j\hbar$.

Bohr elevated this and similar limiting relations into a guiding principle in the old quantum theory. He termed it the Correspondence Principle. Sommerfeld called it a magic wand that only worked in Copenhagen, a back-handed compliment.

The orbital angular momentum number l can grow as large as one wishes. The intrinsic spin, s, cannot— $s = \frac{1}{2}$ for an electron, $s = 1$ for a photon. Even for a large atom like uranium, s would be at most a few hundred, assuming the spin of all the electrons, protons, and neutrons combined constructively. (But such a large value of s would imply an enormous energy which would blow the nucleus apart.) In practice, the classical limit makes no sense for s, another sense in which intrinsic spin is fundamentally non-classical.

The addition rules for angular momentum come from the following considerations: suppose we have two representations $\sigma_1 : SU(2) \to SU(H_1)$ and $\sigma_2 : SU(2) \to SU(H_2)$. H_1 and H_2 are the Hilbert spaces for two separate physical systems. The Hilbert space of the combined system is the tensor product $H_1 \otimes H_2$, and the representation $\sigma_1 \otimes \sigma_2 : SU(2) \to SU(H_1 \otimes H_2)$ begets the angular momentum operator for the combined system. Even if σ_1 and σ_2 are irreducible representations, $\sigma_1 \otimes \sigma_2$ generally won't be, but will decompose into a direct sum of irreducible representations:

$$\sigma_1 \otimes \sigma_2 : SU(2) \to SU(K_1) \oplus \ldots \oplus SU(K_r)$$

The addition rules now all fall out from theorems of group representation theory.

Enough said.