Excursions in Computing Science:
Book 8c. Symmetry: Simplifying Matrices.
Part III Continuous symmetries and the atom.

T. H. Merrett*
McGill University, Montreal, Canada

June 10, 2015

I. Prefatory Notes

1. Which operations preserve the triangle?
2. How did we get the last two reflections?
3. We can abstract all this.
4. Other properties of the operators.
5. Operating on the group operators.
6. ghg^{-1}.
7. The rotation subgroup, $\{(), (123), (132)\}$.
8. Simplifying matrices.
9. Other representations of the group.
10. Simplifying matrices.
11. Regular representation.
12. Molecules
13. Modes of motion of a molecule.
14. Greenhouse gases.
15. The water vapour molecule.
16. Tetrahedron.
17. Hexa/Octahedra.
18. Dodeca/Icosahedra.
19. Infinite groups
20. 1D crystals

21. 2D crystals

22. 2D waves

23. Brillouin zone.

24. Non-translational crystal symmetries.

25. Wallpaper groups.

26. Continuous groups. As we know from triangles, squares, etc., rotations form groups. An n-fold rotation is generated by

$$\begin{pmatrix} \cos\phi_n & -\sin\phi_n \\ \sin\phi_n & \cos\phi_n \end{pmatrix}$$

where $\phi_n = 2\pi/n$. There is no limit to how big the integer n can be.

Arbitrary rotations also form groups. Every element of an arbitrary rotation group is

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

where ϕ is a *parameter* which can take on *any* value between 0 and 2π (with 2π being equivalent to 0).

This is a *continuous* parameter, and the group is a continuous group. "Continuous" is an attribute of mathematics, not of nature. If we believe that nature abhors infinities, it certainly cannot be continuous: continuous math includes not only an (uncountable) infinity of numerical values, but most of them—still an uncountable infinity—require an infinite number of decimal places each.

Finite rotation groups have a smallest angle, namely ϕ_n, and all rotations involve multiples, $1,..,n$, times this.

Since a continuous rotation has no smallest angle, what is the generator of the group?

To see what happens to

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

as ϕ becomes very small, we must introduce the series expansions of cos() and sin().

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + .. \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - .. \end{aligned}$$

You can check this by writing programs to evaluate the series until the next term in the series is smaller than, say, 10^{-6} and display the differences between these series and cos x or sin x, respectively. These calculations must be done in radians: why?

From these we also have

$$e^{ix} = 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \frac{x^6}{6!} - ..$$

and this will be very useful quite soon.

When x is very small, x^2, x^3 and so on can be neglected and

$$\begin{aligned} \cos x &\overset{x\to 0}{\longrightarrow} 1 \\ \sin x &\overset{x\to 0}{\longrightarrow} x \\ \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} &\overset{\phi\to 0}{\longrightarrow} \begin{pmatrix} 1 & -\phi \\ \phi & 1 \end{pmatrix} = I + \phi\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \end{aligned}$$

Now here is something remarkable. *This is all we need.* Let's *define*

$$e^{iX} \stackrel{\text{def}}{=} I + iX - \frac{X^2}{2!} - i\frac{X^3}{3!} + \frac{X^4}{4!} + i\frac{X^5}{5!} - \frac{X^6}{6!} - ..$$

for any *square matrix* X and see what happens when

$$X = -i \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$$

First,

$$\begin{aligned} X^2 &= -\begin{pmatrix} -1 & \\ & -1 \end{pmatrix} = I, \\ X^3 &= X, \\ X^4 &= I, \ \ .. \end{aligned}$$

So

$$\begin{aligned} e^{i\phi X} &= I(1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \frac{\phi^6}{6!} + ..) \\ &\quad + iX(1 - \frac{\phi^3}{3!} + \frac{\phi^5}{5!} - ..) \\ &= I\cos\phi + iX\sin\phi \\ &= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \end{aligned}$$

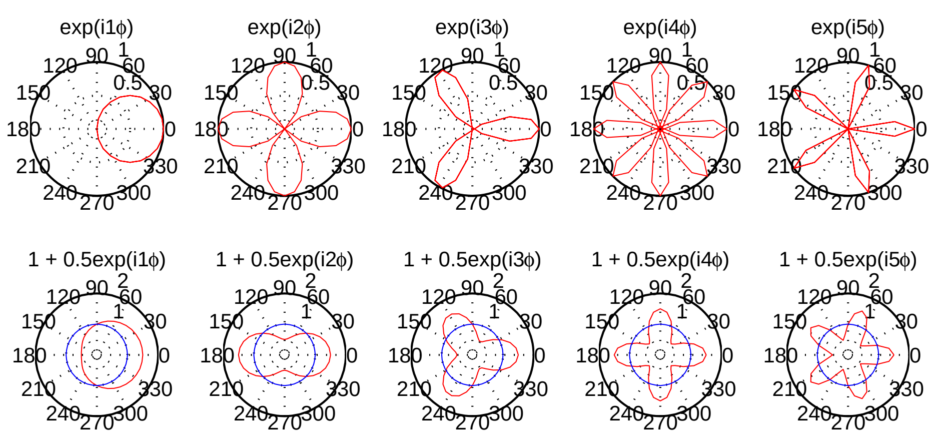
Because we can find any group element, $e^{i\phi X}$, given X, it is appropriate to call X the *generator* of the rotation group. The *parameter* is ϕ.

Since two-dimensional rotations commute, the representations are all one-dimensional.

Since the group has an infinite number of elements (given by the parameter ϕ), there are an infinte number of representations.

The representations are $e^{i\phi}, e^{2i\phi}, e^{3i\phi}, ..$, corresponding to waves with 1 period around the unit circle, 2 periods, 3 periods, and so on.

The figures show both $e^{im\phi}$ and $1 + 0.5e^{im\phi}$ (real parts only) for $m = 1, .., 5$: the latter show the wave nature more explicitly.



These waves would give the amplitudes for quantum-mechanical behaviour under circular symmetry. The probabilities would be the *squares*, i.e., the product of the amplitudes with their 2-number (complex) conjugates. These are all 1, giving equal probabilities in all directions around the circle.

In the expression $e^{im\phi}$, as we just now saw, m is the number of waves that fit around the circle. That is, m is the (circular) wavenumber.

In Week 7a we said that momentum is proportional to wavenumber, k, which appears with x in the 2-number exponential e^{ikx} : $p = \hbar k$.

We will also define *angular momentum*

$$J \stackrel{\text{def}}{=} \hbar m$$

(although we will see in Note 29, where we work in three dimensions, that $\hbar m$ is merely the z-component of the full angular momentum).

(The symbol L is often used for angular momentum, but what we've been calling J is essentially angular momentum so I'll use J for consistency. Often the relationship is $L = \hbar J$, but I hope the presence or absence of the $\hbar$ in different places will not confuse.)

Angular momentum is important because it is also conserved.

Note that, while the angular momentum, $\hbar m$, is known exactly, the angle, ϕ, is not known at all: when we find the probability function, by multiplying the amplitude $e^{i\phi}$ by its 2-number (complex) conjugate, we get 1 everywhere. No angle is preferred. Angle and angular momentum are said to be "complementary" in this way.

27. Spherical symmetry. As the group of circular symmetry has an infinite number of representations, so the group of spherical symmetry will have an infinite number of representations, this time not one-dimensional because 3D rotations do not commute (Week 6 Note 5, Week 7c Note 9).

The rotation groups for discrete subsets of the sphere are suggestive but not conclusive. The rotation subgroup, A4, of the tetrahedron has representations of dimensions 1, 1, 1 and 3 (see Excursions, Book 8cI). The rotation subgroup of the cube and octahedron has representations of dimensions 1, 1, 2, 3 and 3 (see Notes 17 and 16, Book 8cI). The rotation subgroup, A5, of the dodecahedron and icosahedron has representations of dimensions 1, 3, 3, 4 and 5 (see Excursions, Book 8cI).

We have already seen 2- and 3-dimensional representations of three-dimensional rotations in Week 6 Note 3 and Week 7c Note 9, respectively. And there is always the 1-dimensional trivial representation.

We are going to find that the group of spherical symmetry has representations of 1, 2, 3, 4, .. dimensions, with no upper limit.

Let's start by finding the generators of the spherical group. We'll use the 3D representation given by the Interval Algebra in Week 7c Note 9.

The rotation of any 3D vector by an angle α around the axis (p, q, r) (normalized to $p^2+q^2+r^2 = 1$) is given there by

$$\begin{pmatrix} \cos\alpha & -r\sin\alpha & q\sin\alpha \\ r\sin\alpha & \cos\alpha & -p\sin\alpha \\ -q\sin\alpha & p\sin\alpha & \cos\alpha \end{pmatrix} + (1-\cos\alpha)\begin{pmatrix} p \\ q \\ r \end{pmatrix}(p, q, r)$$

For very small α this is approximated, as in the previous Note, by

$$\begin{pmatrix} 1 & -r\alpha & q\alpha \\ r\alpha & 1 & -p\alpha \\ -q\alpha & p\alpha & 1 \end{pmatrix} = I - i\alpha p J'_x - i\alpha q J'_y - i\alpha r J'_z$$

where the *generators* are

$$J'_x = i \begin{pmatrix} & & \\ & & -1 \\ & 1 & \end{pmatrix}$$

$$J'_y = i \begin{pmatrix} & & 1 \\ & & \\ -1 & & \end{pmatrix}$$

$$J'_z = i \begin{pmatrix} & -1 & \\ 1 & & \\ & & \end{pmatrix}$$

(Compare these with the 2D generators, X, in the previous Note.)

We can also write this formally as a dot product involving a vector of matrices, $\vec{J}'$, and a vector af angles, $(\alpha p, \alpha q, \alpha r)$

$$I + i\alpha(p,q,r)\begin{pmatrix} J'_x \\ J'_y \\ J'_z \end{pmatrix} = I + \vec{\alpha}.\vec{J}'$$

The J' do not commute but their *commutators* have interesting properties.

$$[J'_x, J'_y] \stackrel{\text{def}}{=} J'_xJ'_y - J'_yJ'_x = iJ'_z$$

$$[J'_y, J'_z] \stackrel{\text{def}}{=} J'_yJ'_z - J'_zJ'_y = iJ'_x$$

$$[J'_z, J'_x] \stackrel{\text{def}}{=} J'_zJ'_x - J'_xJ'_z = iJ'_y$$

(Compare the Pauli matrices in Week 6 Note 5.)

We can reconstruct the group elements from the generators as formal exponentials, as we did for the circle group elements in the previous Note.

$$e^{i\vec{\alpha}.\vec{J}'} = I + i\vec{\alpha}.\vec{J}' - \frac{(\vec{\alpha}.\vec{J}')^2}{2!} - i\frac{(\vec{\alpha}.\vec{J}')^3}{3!} + i\frac{(\vec{\alpha}.\vec{J}')^4}{4!} +$$

where

$$\vec{\alpha}.\vec{J}' = i\begin{pmatrix} & -r\alpha & q\alpha \\ r\alpha & & -p\alpha \\ -q\alpha & p\alpha & \end{pmatrix} = i\alpha\begin{pmatrix} & -r & q \\ r & & -p \\ -q & p & \end{pmatrix}$$

$$(\vec{\alpha}.\vec{J}')^2 = -\alpha^2\begin{pmatrix} -r^2-q^2 & pq & rp \\ pq & -r^2-p^2 & qr \\ rp & qr & -p^2-q^2 \end{pmatrix}$$

$$= -\alpha^2\begin{pmatrix} p^2-1 & pq & rp \\ pq & q^2-1 & qr \\ rp & qr & r^2-1 \end{pmatrix}$$

$$= \alpha^2(I - \begin{pmatrix} p \\ q \\ r \end{pmatrix}(p,q,r))$$

$$(\vec{\alpha}.\vec{J}')^3 = i\alpha^3\begin{pmatrix} & -r & q \\ r & & -p \\ -q & p & \end{pmatrix} = \alpha^2(\vec{\alpha}.\vec{J}')$$

So

$$
\begin{aligned}
e^{i\vec{\alpha}.\vec{J}'} &= I(1-\frac{\alpha^2}{2!}+\frac{\alpha^4}{4!}-..)-\begin{pmatrix} p \\ q \\ r \end{pmatrix}(p,q,r)(-\frac{\alpha^2}{2!}+\frac{\alpha^4}{4!}-..) \\
&\quad -i\frac{\vec{\alpha}.\vec{J}'}{\mid \alpha \mid}(\alpha-\frac{\alpha^3}{3!}+\frac{\alpha^5}{5!}-..) \\
&= \begin{pmatrix} \cos\alpha & -r\sin\alpha & q\sin\alpha \\ r\sin\alpha & \cos\alpha & -p\sin\alpha \\ -q\sin\alpha & p\sin\alpha & \cos\alpha \end{pmatrix} + (1-\cos\alpha)\begin{pmatrix} p \\ q \\ r \end{pmatrix}(p,q,r)
\end{aligned}
$$

28. Commutator algebra. With the formal definition we've used for exponentials of matrices, one familiar property of exponentials gets lost.

It is no longer true that

$$e^{iX}e^{iY} = e^{i(X+Y)} \qquad \text{[wrong!]}$$

because if X and Y do not commute then e^{iX} and e^{iY} cannot commute.

We need a new result. But it should be apparent that the difference between old and new will involve only the commutator, $[X,Y]$, of X and Y:

$$e^{iX}e^{iY} = e^{i(X+Y)} \times (\text{some function of } [X,Y], \text{ etc.})$$

We can work out some of these $[X,Y]$-dependent terms to make this more convincing.

We will need not only the series expansion for e^{iX} but also of the inverse operation to exponentiation, the *natural logarithm*

$$\ln(1+K) = K - \frac{K^2}{2} + \frac{K^3}{3} - \frac{K^4}{4} + ..$$

Since e^{iX} and e^{iY} are elements of a group (in our case the spherical group) their product must also be a group element, so we can call it e^{iZ}. (Do not confuse X, Y and Z with J'_x, J'_y and J'_z: we are being more general now.)

So we make two definitions for convenience.

$$1+K \stackrel{\text{def}}{=} e^{iZ} \stackrel{\text{def}}{=} e^{iX}e^{iY}$$

so iZ will be $\ln(1+K)$.

Now let's work out the series, going only up to third powers. (The full result requires going to an infinite number of powers, but we'll get tired doing that, and besides our purpose is only to support the general argument that the difference between $e^{iX}e^{iY}$ and $e^{i(X+Y)}$ involves only commutators.)

$$
\begin{aligned}
1+K &= e^{iX}e^{iY} = (1+iX-\frac{X^2}{2!}-\frac{iX^3}{3!}..)(1+iY-\frac{Y^2}{2!}-\frac{iX^3}{3!}..) \\
&= 1+i(X+Y)-\frac{1}{2}(X^2+2XY+Y^2)-\frac{i}{6}(X^3+3X^2Y+3XY^2+Y^3)..
\end{aligned}
$$

So

$$
\begin{aligned}
iZ &= \ln(1+K) = K - \frac{1}{2}K^2 + \frac{1}{3}K^3.. \\
&= i(X+Y)+\frac{1}{2}((X+Y)^2-(X^2+2XY+Y^2)) \\
&\quad +\frac{i}{12}((3(X+Y)(X^2+2XY+Y^2)+3(X^2+2XY+Y^2)(X+Y) \\
&\quad -2(X^3+3X^2Y+3XY^2+Y^3)-4(X+Y)^3).. \\
&= i(X+Y+\frac{i}{2}[X,Y]+\frac{1}{12}([X,[X,Y]]+[Y,[X,Y]])..)
\end{aligned}
$$

Thus Z, to third powers, is $X + Y +$ commutators, as we argued. Clearly it cannot be otherwise when taken to any power.

Now that we have seen that the product of two group elements, g_X and g_Y, having generators X and Y respectively, depends only on X, Y and their commutators.

But in the spherical group of Note 27 the commutators of the generators are also generators: $[J'_x, J'_y] = -iJ'_z$, etc.

We can generalize such relationships to a *commutator algebra*, in which the commutator of any two elements is a linear combination of (some of) the elements, e.g. $[X, Y] = aW + bX + cZ$, and which has the following axioms.

$$
\begin{aligned}
[X+Y, Z] &= [X, Z] + [Y, Z] \\
[aX, Y] &= a[X, Y] \\
[X, Y] &= -[Y, X] \\
0 &= [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]]
\end{aligned}
$$

The first two axioms say that the algebra is linear. The third axiom replaces the definition that $[X, Y] = XY - YX$ (for which it is obviously true) and generalizes the commutator.

29. Representations of the spherical group. We haven't lost sight of the goal of finding the representations of the spherical group.

In Note 27 we saw the 3D representation generated by J'_x, J'_y and J'_z. We also have the 2D representation generated by the half-Pauli matrices $\sigma_x/2$, $\sigma_y/2$ and $\sigma_z/2$ (Week 6 Note 3). And we have the 1D trivial representation, of course.

To explore further, we need to diagonalize. Since the generators do not commute, they cannot all simultaneously be diagonalized.

So we must choose one to diagonalize and take the others as they work out.

The choice of which one is arbitrary. It could equally well be J'_x, J'_y or J'_z. But J'_z is always chosen, and we do the same.

We need Q so that $J_z = QJ'_zQ^{-1}$ is diagonal. Try

$$
Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i & \\ & & -\sqrt{2} \\ -1 & -i & \end{pmatrix} \qquad Q^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & & -1 \\ i & & i \\ & -\sqrt{2} & \end{pmatrix}
$$

Thus, with $J' = pJ'_x + qJ'_y + rJ'_z$

$$
J = QJ'Q^{-1} = \begin{pmatrix} r & \frac{p-iq}{\sqrt{2}} & \\ \frac{p+iq}{\sqrt{2}} & & \frac{p-iq}{\sqrt{2}} \\ & \frac{p+iq}{\sqrt{2}} & -r \end{pmatrix}
$$

So

$$
J_x = \frac{1}{\sqrt{2}}\begin{pmatrix} & 1 & \\ 1 & & 1 \\ & 1 & \end{pmatrix} \quad \text{cf} \quad \frac{\sigma_x}{2} = \frac{1}{2}\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}
$$

$$J_y = \frac{1}{\sqrt{2}}\begin{pmatrix} & -i & \\ i & & -i \\ & i & \end{pmatrix} \quad \text{cf} \quad \frac{\sigma_y}{2} = \frac{1}{2}\begin{pmatrix} & i \\ -i & \end{pmatrix}$$

$$J_z = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix} \quad \text{cf} \quad \frac{\sigma_x}{2} = \begin{pmatrix} 1/2 & \\ & -1/2 \end{pmatrix}$$

where I've also shown the half-Pauli matrices for comparison.

At a guess, in 4D,

$$J_z = \begin{pmatrix} 3/2 & & & \\ & 1/2 & & \\ & & -1/2 & \\ & & & -3/2 \end{pmatrix}$$

But how do we show what J_x and J_y are?

In Week 6 Notes 8 and 9 we constructed a 3D and a 1D representation by taking the tensor product of a 2D representation with itself and symmetrizing. This process can be abbreviated

$$A_{(4)} = Q(A'_{(2)} \otimes A'_{(2)})Q^{-1}$$

$$\text{where} \quad Q = \begin{pmatrix} 1 & & & \\ & 1/\sqrt{2} & 1/\sqrt{2} & \\ & & & 1 \\ & 1/\sqrt{2} & -1/\sqrt{2} & \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 1 & & & \\ & 1/\sqrt{2} & & 1/\sqrt{2} \\ & 1/\sqrt{2} & & -1/\sqrt{2} \\ & & 1 & \end{pmatrix}$$

$$\text{and} \qquad A'_{(2)} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \qquad \text{with } 1 = \det(A'_{(2)}) = ad - bc$$

We got

$$A_{(4)} = \begin{pmatrix} a^2 & ac\sqrt{2} & c^2 & \\ ab\sqrt{2} & ad+bc & cd\sqrt{2} & \\ b^2 & bd\sqrt{2} & d^2 & \\ & & & 1 \end{pmatrix} = A'_{(3)} \oplus A_{(1)}$$

We can similarly build a 4D and a 2D representation from the tensor product of $A'_{(2)}$ and $A'_{(3)}$

$$A_{(6)} = Q(A'_{(2)} \otimes A'_{(3)})Q^{-1}$$

where

$$Q = \begin{pmatrix} 1 & & & & & \\ & \sqrt{2/3} & & \sqrt{1/3} & & \\ & & \sqrt{1/3} & & \sqrt{2/3} & \\ & & & & & 1 \\ & -\sqrt{1/3} & & \sqrt{2/3} & & \\ & & -\sqrt{2/3} & & \sqrt{1/3} & \end{pmatrix}$$

The result is

$$\begin{aligned} A_{(6)} &= \begin{pmatrix} a^3 & a^2c\sqrt{3} & ac^2\sqrt{3} & c^3 & & \\ a^2b\sqrt{3} & a+abc & c+acd & c^2d\sqrt{3} & & \\ ab^2\sqrt{3} & b+abd & d+bcd & cd^2\sqrt{3} & & \\ b^3 & b^2d\sqrt{3} & bd^2\sqrt{3} & d^3 & & \\ & & & & a & c \\ & & & & b & d \end{pmatrix} \\ &= A'_{(4)} \oplus A'_{(2)} \end{aligned}$$

You can check these calculations, but the point of making them is to motivate the 4-dimensional J matrices.

In symmetrizing $A'_{(2)} \otimes A'_{(2)}$, Q contains some $1/\sqrt{2}$s, which J_x and J_y also contain. They also came from normalizing the symmetric $\frac{1}{\sqrt{2}}(| +- > + | -+ >)$, and antisymmetric, $\frac{1}{\sqrt{2}}(| +- > - | -+ >)$, state combinations.

The $\sqrt{2/3}$ and $\sqrt{1/3}$ in Q for $A'_{(2)} \otimes A'_{(3)}$ come from the 4D J_x and J_y. We'll now see how.

To get the results quickly, I'm going to make an argument which is not watertight. My justification will be that it gives the coefficients in Q which block-diagonalized $A'_{(2)} \otimes A'_{(3)}$.

In Week 6 we associated the 2D representation with "spin 1/2", a situation which must be rotated through 4π to restore it to itself—rotating by 2π changes the sign of the amplitude. We also associated the 3D representation with "spin 1", a combination of two spin-1/2s.

So the dimension of the representation is $2\ell + 1$ where ℓ is the "spin"—we'll call it "anguℓar momentum" for this discussion. Inversely the angular momentum relates to the dimension, d, of the representation: $\ell = (d-1)/2$.

The (diagonal) elements of J_z are $\ell, \ell-1, .., -\ell : d = 2\ell+1$ of them. This I am going to suppose is true for all dimensions.

Given that, we can find J_x and J_y for any dimension by a trick.

The trick is to consider the "raising operator" $J_+ \stackrel{\text{def}}{=} (J_x + iJ_y)/\sqrt{2}$ and the "lowering operator" $J_- \stackrel{\text{def}}{=} (J_x - iJ_y)/\sqrt{2}$. (The $1/\sqrt{2}$ is not always used, but it normalizes the operators.)

From the general commutator relationships, $[J_x, J_y] = iJ_z$, etc.

$$[J_+ . J_-] = J_z$$

as well as

$$[J_\pm, J_z] = \pm J_\pm$$

Thus we have

$$\begin{aligned}
\text{In 2D}\ J_+ &= \frac{1}{\sqrt{2}}\frac{1}{2}(\sigma_x + i\sigma_y) = \frac{1}{\sqrt{2}}\begin{pmatrix} & 1 \\ & \end{pmatrix} \\
J_- &= \frac{1}{\sqrt{2}}\frac{1}{2}(\sigma_x - i\sigma_y) = \frac{1}{\sqrt{2}}\begin{pmatrix} & \\ 1 & \end{pmatrix} \\
\text{In 3D}\ J_+ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(J_x + iJ_y) = \frac{1}{\sqrt{2}}\begin{pmatrix} & 1 & \\ & & 1 \\ & & \end{pmatrix} \\
J_- &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(J_x - iJ_y) = \frac{1}{\sqrt{2}}\begin{pmatrix} & & \\ 1 & & \\ & 1 & \end{pmatrix} \\
\text{In 4D}\ J_+ &= \begin{pmatrix} & c_{3/2} & & \\ & & c_{1/2} & \\ & & & c_{-1/2} \\ & & & \end{pmatrix} \quad J_x = \frac{1}{\sqrt{2}}\begin{pmatrix} & c_{3/2} & & \\ c_{3/2} & & c_{1/2} & \\ & c_{1/2} & & c_{-1/2} \\ & & c_{-1/2} & \end{pmatrix} \\
J_- &= \begin{pmatrix} & & & \\ c_{3/2} & & & \\ & c_{1/2} & & \\ & & c_{-1/2} & \end{pmatrix} \quad J_y = \frac{i}{\sqrt{2}}\begin{pmatrix} & -c_{3/2} & & \\ c_{3/2} & & -c_{1/2} & \\ & c_{1/2} & & -c_{-1/2} \\ & & c_{-1/2} & \end{pmatrix}
\end{aligned}$$

where we don't yet know the cs.

But we can find them out with $[J_+, J_-] = J_z$.

$$J_+J_- = \begin{pmatrix} & c_\ell & & \\ & & c_{\ell-1} & \\ & & & \ddots \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} & & & \\ c_\ell & & & \\ & c_{\ell-1} & & \\ & & \ddots & \end{pmatrix} = \begin{pmatrix} c_\ell^2 & & & \\ & c_{\ell-1}^2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix}$$

$$J_-J_+ = \begin{pmatrix} & & & \\ c_\ell & & & \\ & c_{\ell-1} & & \\ & & \ddots & \end{pmatrix} \begin{pmatrix} & c_\ell & & \\ & & c_{\ell-1} & \\ & & & \ddots \\ & & & & \ddots \end{pmatrix} = \begin{pmatrix} 0 & & & \\ & c_\ell^2 & & \\ & & c_{\ell-1}^2 & \\ & & & \ddots \end{pmatrix}$$

So

$$\begin{pmatrix} c_\ell^2 & & & \\ & c_\ell^2 - c_{\ell-1}^2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix} = [J_+, J_-] = J_z = \begin{pmatrix} \ell & & & \\ & \ell - 1 & & \\ & & \ddots & \\ & & & -\ell \end{pmatrix}$$

That is

$$\begin{aligned} c_\ell^2 &= \ell \\ c_{\ell-1}^2 &= \ell + \ell - 1 = 2\ell - 1 \\ c_{\ell-2}^2 &= \ell + \ell - 1 + \ell - 2 = 3\ell - 3 \\ c_m^2 = c_{\ell-k}^2 &= \sum_{j=0}^{k} (\ell - j) = (k+1)\ell - (k+1)k/2 \\ &= (k+1)(2\ell - k)/2 \\ &= (\ell - m + 1)(\ell + m)/2 \end{aligned}$$

So

$$c_m = \sqrt{(\ell - m + 1)(\ell + m)/2}$$

giving

$$\begin{array}{llll} \text{for 2D} & \ell & = 1/2 & c_{1/2} = 1/\sqrt{2} \\ \text{for 3D} & \ell & = 1 & c_1 = 1 \\ & & & c_0 = 1 \\ \text{for 4D} & \ell & = 3/2 & c_{3/2} = \sqrt{3/2} \\ & & & c_{1/2} = \sqrt{2} \\ & & & c_{-1/2} = \sqrt{3/2} \end{array}$$

and so on. Note that $c_m = c_{1-m}$.

This general result agrees with what we knew for 2D and 3D and with what we guessed for 4D.

Back to 4D

$$J_+ = \begin{pmatrix} & \sqrt{\frac{3}{2}} & & \\ & & \sqrt{2} & \\ & & & \sqrt{\frac{3}{2}} \\ & & & \end{pmatrix} \qquad J_- = \begin{pmatrix} & & & \\ \sqrt{\frac{3}{2}} & & & \\ & \sqrt{2} & & \\ & & \sqrt{\frac{3}{2}} & \end{pmatrix}$$

$$J_x = \frac{1}{\sqrt{2}} \begin{pmatrix} & \sqrt{\frac{3}{2}} & & \\ \sqrt{\frac{3}{2}} & & \sqrt{2} & \\ & \sqrt{2} & & \sqrt{\frac{3}{2}} \\ & & \sqrt{\frac{3}{2}} & \end{pmatrix} \quad J_y = \frac{i}{\sqrt{2}} \begin{pmatrix} & -\sqrt{\frac{3}{2}} & & \\ \sqrt{\frac{3}{2}} & & -\sqrt{2} & \\ & \sqrt{2} & & -\sqrt{\frac{3}{2}} \\ & & \sqrt{\frac{3}{2}} & \end{pmatrix}$$

Let's confirm the Q we used to block diagonalize $A'_{(2)} \otimes A'_{(3)}$ to $A'_{(4)} \oplus A'_{(2)}$. The coefficient $\sqrt{3/2}$ has appeared in $J_\pm$, J_x and J_y but not the $\sqrt{2/3}$ or $\sqrt{1/3}$ that we used.

We use the notation $\mid \ell, m >$ to specify the m^{th} basis vector in representation ℓ of $2\ell+1$ dimensions $(-\ell \le m \le \ell)$. To specify the tensor product of two representations, we combine the basis vectors.

Thus, the $m = 3/2$ state of the 4D representation $\ell = 3/2$ can be assembled from an $m = 1/2$ state of the 2D representation $\ell = 1/2$ and an $m = 1$ state of the 3D representation $\ell = 1$ in only one way.

$$\mid \frac{3}{2}, \frac{3}{2} > = \mid \frac{1}{2}, \frac{1}{2} > \mid 1, 1 >$$

The $m = \ell - 1$ states are obtained from this by using the lowering operator J_-

$$J_- \mid \frac{3}{2}, \frac{3}{2} > = J_-(\mid \frac{1}{2}, \frac{1}{2} > \mid 1, 1 >) = (J_- \mid \frac{1}{2}, \frac{1}{2} >) \mid 1, 1 > + \mid \frac{1}{2}, \frac{1}{2} > J_-(\mid 1, 1 >)$$

Consult the J_- matrices for 4D ($\ell = 3/2$), 3D ($\ell = 1$) and 2D ($\ell = 1/2$):

$$\begin{aligned} J_- \mid \frac{3}{2}, \frac{3}{2} > &= \sqrt{\frac{3}{2}} \mid \frac{3}{2}, \frac{1}{2} > \\ J_- \mid \frac{1}{2}, \frac{1}{2} > &= \sqrt{\frac{1}{2}} \mid \frac{1}{2}, -\frac{1}{2} > \\ J_- \mid 1, 1 > &= \mid 1, 0 > \end{aligned}$$

So

$$\sqrt{\frac{3}{2}} \mid \frac{3}{2}, \frac{1}{2} > = \sqrt{\frac{1}{2}} \mid \frac{1}{2}, -\frac{1}{2} > \mid 1, 1 > \; + \mid \frac{1}{2}, \frac{1}{2} > \mid 1, 0 >$$

Hence

$$\mid \frac{3}{2}, \frac{1}{2} > = \sqrt{\frac{1}{3}} \mid \frac{1}{2}, -\frac{1}{2} > \mid 1, 1 > \; + \sqrt{\frac{2}{3}} \mid \frac{1}{2}, \frac{1}{2} > \mid 1, 0 >$$

Similarly

$$\begin{aligned} \sqrt{2} \mid \frac{3}{2}, -\frac{1}{2} > &= J_- \mid \frac{3}{2}, \frac{1}{2} > = J_-(\sqrt{\frac{1}{3}} \mid \frac{1}{2}, -\frac{1}{2} > \mid 1, 1 > + \sqrt{\frac{2}{3}} \mid \frac{1}{2}, \frac{1}{2} > \mid 1, 0 >) \\ &= \sqrt{\frac{1}{3}} (0 + \mid \frac{1}{2}, -\frac{1}{2} > \mid 1, 0 >) + \sqrt{\frac{2}{3}} (\mid \frac{1}{2}, \frac{1}{2} > \mid 1, 0 > + \mid \frac{1}{2}, \frac{1}{2} > \mid 1, -1 >) \\ &= 2\sqrt{\frac{1}{3}} \mid \frac{1}{2}, -\frac{1}{2} > \mid 1, 0 > + \sqrt{\frac{2}{3}} \mid \frac{1}{2}, \frac{1}{2} > \mid 1, -1 > \end{aligned}$$

So

$$\mid \frac{3}{2}, -\frac{1}{2} >= \sqrt{\frac{2}{3}} \mid \frac{1}{2}, -\frac{1}{2} >\mid 1, 0 > + \sqrt{\frac{1}{3}} \mid \frac{1}{2}, \frac{1}{2} >) \mid 1, -1 >$$

Finally you can show similarly that

$$\mid \frac{3}{2}, -\frac{3}{2} >=\mid \frac{1}{2}, -\frac{1}{2} >\mid 1, -1 >$$

These four states of $A'_{(4)}$ and the coefficients give the first four rows of Q.

For the $A'_{(2)}$ we start with $\mid \frac{1}{2}, \frac{1}{2} >$ which can only be some linear combination of $\mid \frac{1}{2}, -\frac{1}{2} >\mid 1, 1 >$ and $\mid \frac{1}{2}, \frac{1}{2} >\mid 1, 0 >$.

This must be orthogonal to $\sqrt{\frac{1}{3}} \mid \frac{1}{2}, -\frac{1}{2} >\mid 1, 1 > + \sqrt{\frac{2}{3}} \mid \frac{1}{2}, \frac{1}{2} >\mid 1, 0 >$, the linear combination that gave us $\mid \frac{3}{2}, \frac{1}{2} >$, so we must have

$$\mid \frac{1}{2}, \frac{1}{2} > = \sqrt{\frac{2}{3}} \mid \frac{1}{2}, -\frac{1}{2} >\mid 1, 1 > - \sqrt{\frac{1}{3}} \mid \frac{1}{2}, \frac{1}{2} >\mid 1, 0 >$$

Applying J_- to both sides gives

$$\mid \frac{1}{2}, -\frac{1}{2} > = \sqrt{\frac{1}{3}} \mid \frac{1}{2}, -\frac{1}{2} >\mid 1, 0 > - \sqrt{\frac{2}{3}} \mid \frac{1}{2}, \frac{1}{2} >\mid 1, -1 >$$

and this completes Q.

Summarizing this long Note, we have found that the representations of the spherical group are of dimension $d = 1, 2, 3, 4, ..$ They are characterized by an "angular momentum" $\ell = (d-1)/2$, i.e., $\ell = 0,\ 1/2,\ 1,\ 3/2,\ ..$ Each basis vector within a representation ℓ is characterized by a "magnetic quantum number" $m = \ell,\ \ell - 1,\ ..,\ -\ell$. These are the eigenvalues of J_z.

It is also true that $J^2 = J_x^2 + J_y^2 + J_z^2 = \ell(\ell+1)I$ which commutes with each of J_x, J_y and J_z.

All this can be interpreted as describing a total angular momentum vector of magnitude $\sqrt{\ell(\ell+1)}$, oriented in $2\ell+1$ discrete directions that have projections $\ell,\ \ell-1,\ ..,\ -\ell$ on the z-axis. (Remember we arbitrarily chose axes in the representation so that J_z is diagonal.)

The commuting quantities are given by simultaneously diagonalizable operators (matrices) and may be measured simultaneously: thus J^2 and J_z, but alternatively J^2 and J_x or J^2 and J_y.

The discussion of tensor products can be interpreted as the addition of angular momenta. In the first example, $A'_{(2)} \otimes A'_{(2)}$ combined $\ell_1 = 1/2$ with $\ell_2 = 1/2$ to give two outcomes $\ell = \ell_1 + \ell_2 = 1$ and $\ell = \ell_1 - \ell_2 = 0$. For each of these, $m = \ell,\ \ell-1,\ ..,\ -\ell$, i.e., $-1,\ 0,\ 1$ for $\ell = 1$ and 0 for $\ell = 0$.

In the second example, $A'_{(2)} \otimes A'_{(3)}$ combined $\ell_1 = 1/2$ with $\ell_2 = 1$ giving $\ell = \ell_1 - \ell_2 = 3/2$ (m is $-3/2,\ -1/2,\ 1/2,\ 3/2$) and $\ell = \ell_1 - \ell_2$ (m is $-1/2,\ 1/2$).

In general ℓ_1 and ℓ_2 combine to give

$$(2\ell_1 + 1)(2\ell_2 + 1) \quad = \quad 2(\ell_1 + \ell_2) + 1 + 2(\ell_1 + \ell_2 - 1) + 1 + .. + 2 \mid \ell_1 - \ell_2 \mid +1$$

states, consisting of $n = 2\ \min(\ell_1, \ell_2) + 1$ groups of states of angular momentum $\ell = 2(\ell_1 + \ell_2 - k) + 1$ for $k = 0, .., \min(\ell_1, \ell_2)$, each of which has $2\ell + 1$ different states $m = -\ell, .., \ell$.

It is interesting to show the sums as a table, in which $M_j = 2\ell_j + 1$ and the entries are

M_1M_2, #
decomposition

where the decomposition has $\# = 2\ \min(\ell_1, \ell_2)$ terms and sums to M_1M_2.

ℓ_2	$\ell_1 =$ $M_1 =$ M_2	1/2 2	1 3	3/2 4	2 5	5/2 6
$\frac{1}{2}$	2	4,2 3+1	6,2 4+2	8,2 5+3	10,2	12,2
1	3		9,3 5+3+1	12,3 6+4+2	15,3	18,3
$\frac{3}{2}$	4			16,4 7+5+3+1	20,4 8+6+4+2	24,4
2	5				25,5 9+7+5+3+1	30,5 10+8+6+4+2

30. Spherical harmonics. The culmination of our work on the simpler finite groups, and on the circle group, was to show the possible modes of "vibration", or the general physical states, permitted by the symmetry. To do this with continuous groups we may no longer restrict our attention to a finite number of points but must consider an infinity—indeed a continuous infinity—of points which can "move".

Since we claim nature cannot be continuous it is best to think of the following discussion as giving rules for calculating the "vibration" of any one point we happen to take an interest in.

But these rules require us to look at continuous math, whether it describes continuous nature or is merely an approximation to discrete nature. Continuous math, particularly the calculation of slopes, has been developed to a high degree and is worth using even if only as a model.

We wrote a solution for the vibrations in circle symmetry, namely $e^{im\phi}$. We also found a diagonal matrix for the m states in spherical symmetry, namely

$$J_z = \begin{pmatrix} m & & & \\ & m-1 & & \\ & & \ddots & \\ & & & -m \end{pmatrix}$$

What J_z does is give the eigenvalue m or $m-1$ or .. or $-m$ when applied to a vibration state characterized by the quantum number

$$J_z \mid m > = m \mid m >$$

Let's put circle and z-part of sphere together and suppose $\mid m > = e^{im\phi}$

$$J_z e^{im\phi} = m e^{im\phi}$$

We found out (Excursion to Note 26 on the series for $\cos(x)$, $\sin(x)$ and e^{ix}) that

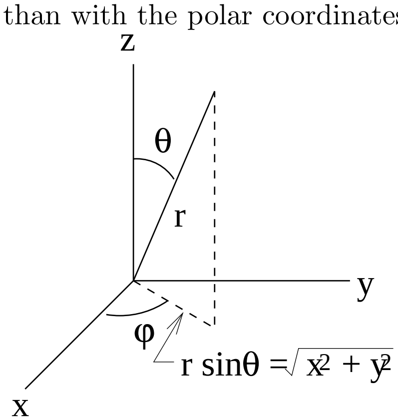
$$\text{Slope}_\phi e^{im\phi} = im e^{im\phi}$$

This makes it plausible to identify the two operators

$$J_z = -i\ \text{Slope}_\phi$$

One is a matrix, the other the slope operator of continuous math.

Now let's change some variables, since we have been working with J_x, J_y, J_z and even $(J_x \pm iJ_y)/\sqrt{2}$, rather than with the polar coordinates r, θ and ϕ.



$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \\ \text{slope}_\phi x &= -r\sin\theta\sin\phi = -y \\ \text{slope}_\phi y &= -r\sin\theta\cos\phi = x \end{aligned}$$

From these slopes we can work out the effect of slope_ϕ on any function $f(x,y)$.

But we should first summarize slopes for handy reference.

Slope Rules

$\text{slope}_x x^n = nx^{n-1}$ any n	power rule
$\text{slope}_x \sin x = \cos x$ $\text{slope}_x \cos x = -\sin x$	trig rules
$\text{slope}_x f(x)g(x) = (\text{slope}_x f)g + f(\text{slope}_x g)$	product rule
$\text{slope}_x f(g(x)) = \text{slope}_g f \ \text{slope}_x g$ $\text{Slope}_x f(p,q) = \text{slope}_p f \ \text{slope}_x p +\ \text{slope}_q f \text{slope}_x q$	chain rules
$\text{slope}_x f^{-1}(x) = 1/\text{slope}_x f(x)$ $\text{slope}_x \cos^{-1} x = -1/\sqrt{1-x^2}$ $\text{slope}_x \tan^{-1} x = 1/(1+x^2)$	inverse rule inverse trig. rules

Armed with these rules (and this is a good exercise in using the rules, too) we can show several useful things. First

$$\begin{aligned} \text{Slope}_\phi f(x,y) &= \text{slope}_x f \ \text{slope}_\phi x + \text{slope}_y f \ \text{slope}_\phi y \\ &= x \ \text{slope}_y f - y \ \text{slope}_x f \\ &= (x \ \text{slope}_y - y \ \text{slope}_x) f(x,y) \end{aligned}$$

From this we can identify the *operators*

$$\text{Slope}_\phi = x \ \text{slope}_y - y \ \text{slope}_x$$

Therefore $J_z = -i \ \text{Slope}_\phi = -i(x \ \text{slope}_y - y \ \text{slope}_x)$.

Let's make a leap and suppose

$$J_x = -i(y \ \text{slope}_z - z \ \text{slope}_y)$$

and

$$J_y = -i(z \ \text{slope}_x - x \ \text{slope}_z)$$

With a bit of careful slope-finding, we can show

$$[J_x, J_y] = x \ \text{slope}_y - y \ \text{slope}_x = iJ_z$$

which is the correct commutator, so the leap seems to be justified.

We will need the raising and lowering operators

$$\begin{aligned}
\frac{1}{\sqrt{2}}(J_x + iJ_y) &= \frac{1}{\sqrt{2}}(-(x+iy)\text{slope}_z + z(\text{slope}_x + i\ \text{slope}_y)) \\
\frac{1}{\sqrt{2}}(J_x - iJ_y) &= \frac{1}{\sqrt{2}}((x-iy)\text{slope}_z - z(\text{slope}_x - i\ \text{slope}_y))
\end{aligned}$$

The second set of applications of the slope rules is to go back again to polar coordinates, this time to express $(J_x \pm iJ_y)/\sqrt{2}$ as slope operators on θ and ϕ.

From the diagram we had x, y and z in terms of θ and ϕ. Now we need slope$_x$, slope$_y$ and slope$_z$.

$$\theta = \cos^{-1}\tfrac{z}{r} \qquad\qquad \phi = \tan^{-1}\tfrac{y}{x}$$

Therefore

$$\begin{aligned}
\text{slope}_x\theta &= 0 = \text{slope}_y\theta \\
\text{slope}_z\theta &= \text{slope}_{z/r}\theta\ \text{slope}_z\frac{z}{r} = \frac{-1}{\sqrt{x^2+y^2}} = \frac{-1}{r\sin\theta} \\
\text{slope}_z\phi &= 0 \\
\text{slope}_y\phi &= \text{slope}_{y/x}\phi\ \text{slope}_y\frac{y}{x} = \frac{x}{x^2+y^2} = \frac{\cos\phi}{\sqrt{x^2+y^2}} = \frac{\cos\phi}{r\sin\theta} \\
\text{slope}_x\phi &= \text{slope}_{y/x}\phi\ \text{slope}_x\frac{y}{x} = \frac{-y}{x^2+y^2} = \frac{-\sin\phi}{\sqrt{x^2+y^2}} = \frac{-\sin\phi}{r\sin\theta}
\end{aligned}$$

And so

$$\begin{aligned}
\text{slope}_x f(\theta,\phi) &= \text{slope}_\theta f\ \text{slope}_x\theta + \text{slope}_\phi f\ \text{slope}_x\phi \\
&= \frac{-\sin\phi}{r\sin\theta}\ \text{slope}_\phi f(\theta,\phi) \\
\text{slope}_y f(\theta,\phi) &= \text{slope}_\theta f\ \text{slope}_y\theta + \text{slope}_\phi f\ \text{slope}_y\phi \\
&= \frac{\cos\phi}{r\sin\theta}\ \text{slope}_\phi f(\theta,\phi) \\
\text{slope}_z f(\theta,\phi) &= \text{slope}_\theta f\ \text{slope}_z\theta + \text{slope}_\phi f\ \text{slope}_z\phi \\
&= \frac{-1}{r\sin\theta}\ \text{slope}_\theta f(\theta,\phi)
\end{aligned}$$

Now we can write the raising and lowering operators in term of θ and ϕ.

$$\begin{aligned}
J_+ = \frac{1}{\sqrt{2}}(J_x + iJ_y) &= \frac{1}{\sqrt{2}}(-r\sin\theta e^{i\phi}\left(\frac{-1}{r\sin\theta}\right)\text{slope}_\theta + \frac{r\cos\theta}{r\sin\theta}(-\sin\phi + i\cos\phi)\ \text{slope}_\phi) \\
&= \frac{e^{i\phi}}{\sqrt{2}}(\text{slope}_\theta + i\cot\theta\ \text{slope}_\phi) \\
J_- = \frac{1}{\sqrt{2}}(J_x - iJ_y) &= \frac{1}{\sqrt{2}}(r\sin\theta e^{-i\phi}\left(\frac{-1}{r\sin\theta}\right)\text{slope}_\theta - \frac{r\cos\theta}{r\sin\theta}(-\sin\phi - i\cos\phi)\ \text{slope}_\phi) \\
&= -\frac{e^{-i\phi}}{\sqrt{2}}(\text{slope}_\theta - i\cot\theta\ \text{slope}_\phi)
\end{aligned}$$

Now that we have $J_\pm$ in terms of the two angular coordinates, θ and ϕ, on the surface of the sphere, we can find out the "vibrations" on that surface.

We will call the vibration

$$Y_{\ell m}(\theta, \phi) = \Theta_{\ell m}(\theta)\frac{e^{im\phi}}{\sqrt{2\pi}}$$

where the $1/\sqrt{2\pi}$ is a normalizing constant so that $\frac{e^{im\phi}}{\sqrt{2\pi}}\frac{e^{-im\phi}}{\sqrt{2\pi}}$ sums to 1 over the whole circle $\phi = 0..2\pi$.

$Y_{\ell m}(\theta, \phi)$ are called the "spherical harmonics".

$\Theta_{\ell m}(\theta)$ contains the θ component ($\theta = 0..\pi$) and depends on the two quantum numbers ℓ and m that we found in Note 29.

We use J_+ to take advantage of what we know, that $m \le \ell$:

$$J_+ Y_{\ell\ell}(\theta, \phi) = 0$$

This will give us $\Theta_{\ell\ell}(\theta)$. Then we can use J_- repeatedly to give us $\Theta_{\ell m}(\theta)$ for $m < \ell$:

$$J_- Y_{\ell m}(\theta, \phi) = \sqrt{(\ell - m + 1)(\ell + m)/2}Y_{\ell m-1}(\theta, \phi)$$

where the $\sqrt{(\ell - m + 1)(\ell + m)/2}$ is just the element, c_m, of the matrices J_- (and J_+) in Note 29.

Here is the calculation with J_+

$$\begin{aligned} 0 &= \frac{e^{i\phi}}{\sqrt{2}}(\text{slope}_\theta + i\cot\theta\ \text{slope}_\phi)\Theta_{\ell\ell}\frac{e^{im\phi}}{\sqrt{2\pi}} \\ &= \frac{e^{i(m+1)\phi}}{\sqrt{4\pi}}(\text{slope}_\theta\Theta_{\ell\ell}(\theta) - \ell\cot\theta\ \Theta_{\ell\ell}(\theta)) \end{aligned}$$

which is satisfied by

$$\Theta_{\ell\ell} = (-1)^\ell\frac{\sqrt{(2\ell+1)/2}}{2^\ell \ell!}\sin^\ell\theta$$

Check it out with $\Theta_{\ell\ell} = \sin^\ell\theta$: the constant in front is for normalization. We won't derive it, and we could forget it with almost no harm for our purposes, but with it our results will agree with the tables of spherical harmonics that you might look up.)

We can use this to calculate the first few $\ell\ell$ spherical harmonics.

$$\begin{aligned} Y_{00} &= \frac{1}{\sqrt{4\pi}} \\ Y_{11} &= -\frac{1}{4}\sqrt{\frac{6}{\pi}}\sin\theta e^{i\phi} \\ Y_{22} &= \frac{1}{8}\sqrt{\frac{30}{\pi}}\sin^2\theta e^{2i\phi} \\ Y_{33} &= -\frac{1}{8}\sqrt{\frac{35}{\pi}}\sin^3\theta e^{3i\phi} \end{aligned}$$

Now we can find the others, just by applying J_-

$$Y_{\ell,\ell-1} = \frac{J_-}{\sqrt{\ell}}Y_{\ell\ell}$$

e.g.,

$$\begin{aligned} Y_{10} &= -\frac{e^{-i\phi}}{\sqrt{2}}(\text{slope}_\theta - i\cot\theta\,\text{slope}_\phi)(-\frac{1}{4}\sqrt{\frac{6}{\pi}}\sin\theta e^{i\phi}) \\ &= \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta \end{aligned}$$

$$Y_{\ell,\ell-2} = \frac{J_-}{\sqrt{2\ell-1}}Y_{\ell,\ell-1}$$

e.g.,

$$\begin{aligned} Y_{1,-1} &= -\frac{e^{-i\phi}}{\sqrt{2}}(\text{slope}_\theta - i\cot\theta\,\text{slope}_\phi)(\frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta) \\ &= \frac{1}{4}\sqrt{\frac{6}{\pi}}\sin\theta e^{-i\phi} \end{aligned}$$

Since this work has already been published in many tables of spherical harmonics we won't derive any more.

But we can make pictures. Here are some, plotted both as a Mercator projection on the θ-ϕ plane and in colours on the sphere.

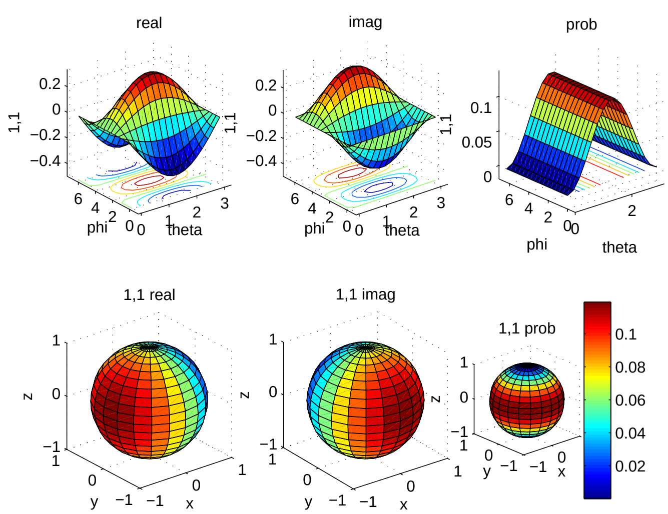
The plots show the "nodes", a word which in discussions of vibrations means the parts of the sphere that never move: the spherical harmonic has value 0 on these lines separating the two colours yellow-green and turquoise in the real and imaginary plots.

Note that there are exactly ℓ such lines.

Note the "dumbell" shapes of $Y_{2,m}(M \neq 0)$.

Note the "tetrahedral" shapes of $Y_{3,\pm2}$.

Here is $Y_{11}(\theta,\phi)$ $(\ell, m = 1)$



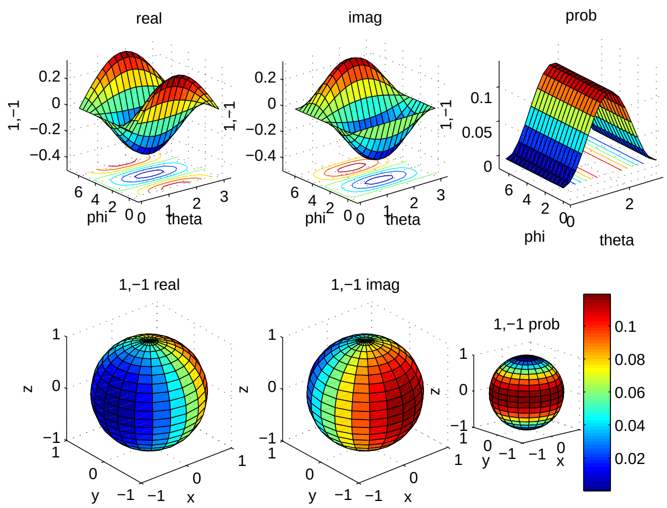
The imaginary part is the real part rotated through a right angle.

There is one nodal line, best seen in the θ-ϕ diagram of the imaginary part, but also evident in planes bisecting the real and imaginary spheres and including the poles.

The probability is independent of ϕ.

The probability pictures show things all happening at the equator.

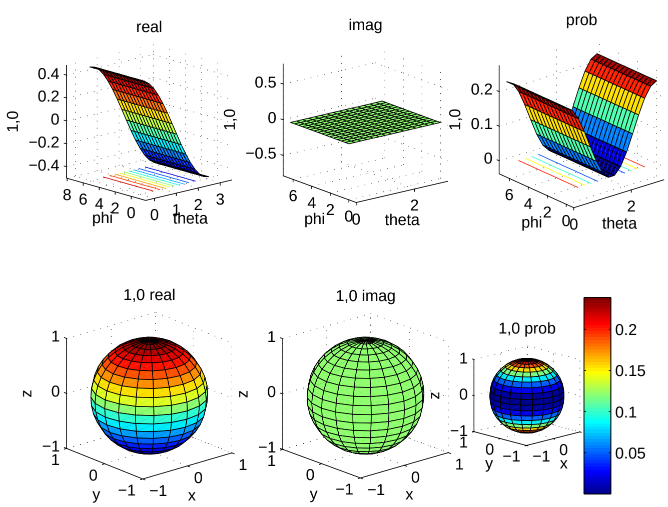
Here is $Y_{1,-1}(\theta, \phi)$ $(\ell, m = 1, -1)$



The real part is the real part of Y_{11} rotated through two right angles. The imaginary part is the same as that of Y_{11}.

There is one nodal line, as before.

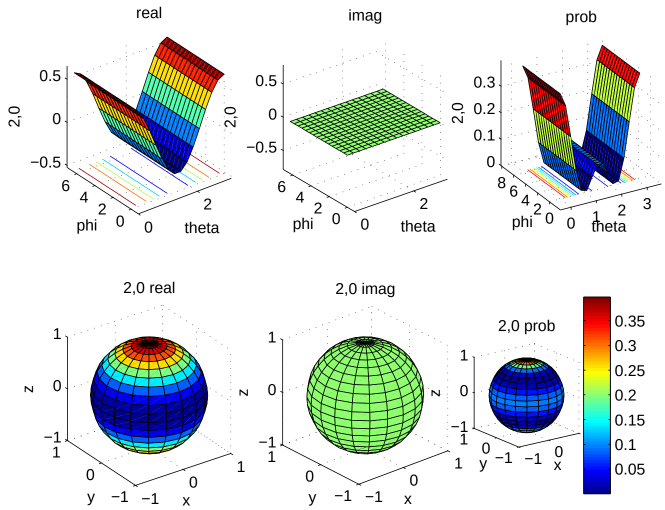
Here is $Y_{10}(\theta, \phi)$ $(\ell, m = 1, 0)$



The imaginary part is 0 everywhere.

The probability pictures show things all happening at the poles.

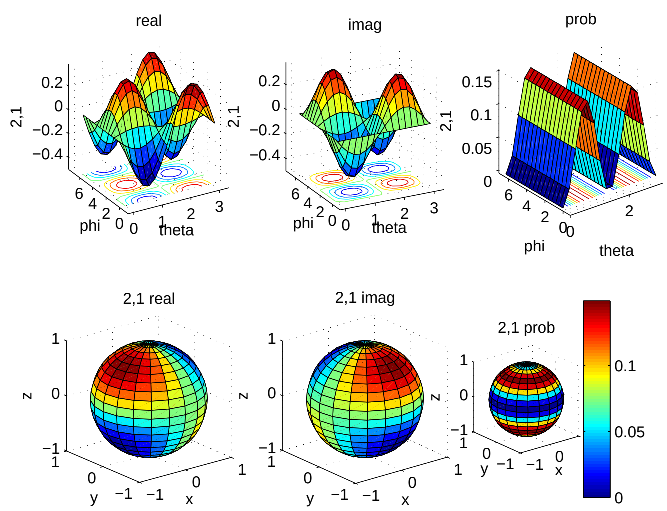
Here is $Y_{20}(\theta, \phi)$ $(\ell, m = 2, 0)$



There are two nodal lines; between the equator and each pole.

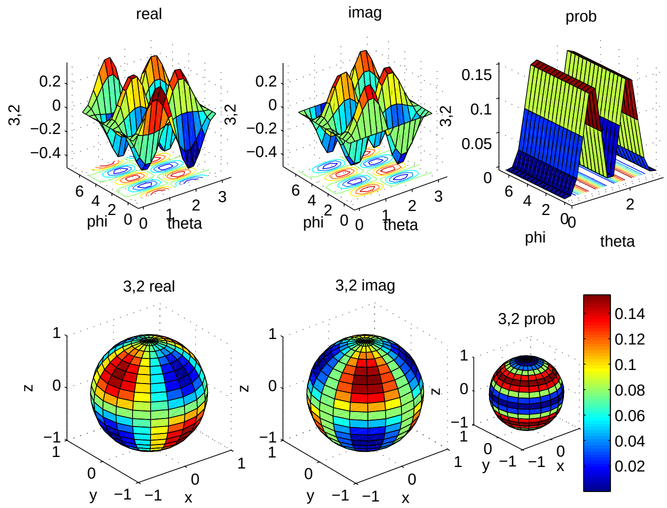
The probabilities are highest at the poles, as with Y_{10}, but there is a small probability of things happening around the equator.

Here is $Y_{21}(\theta, \phi)$ $(\ell, m = 2, 1)$



There are two nodal lines, best seen in the imaginary θ-ϕ plot as a line at $\theta = \pi/2$ crossing a line at $\phi = \pi$. These can also be seen as two planes at right angles to each other and including the poles in either the real or imaginary spheres.

Here is $Y_{32}(\theta,\phi)$ $(\ell, m = 3, 2)$



There are three nodal lines: those of Y_{21} and the equator.

Note the tetrahedra of high amplitudes and of low amplitudes in the real and imaginary spheres.

If these were describing quantum systems, we would *observe* only the probabilities. But the amplitudes are what *interact* with other quantum systems, and so the geometry of both the real and imaginary parts is important.

In the next Note we will see that knowing the radial behaviour, that is, with r as well as with θ and ϕ, helps visualization. We will also see that adding the radial behaviour to the spherical harmonics describes atoms.

Trying to understand the atom drove the creation of quantum mechanics in the first three decades of the 20th century.

31. Atomic physics. The spherical symmetry we have been discussing describes, among other things, the electron structure of the atom.

Shortly after he left McGill, Ernest Rutherford discovered that the atom is made up of electrons orbiting a very small, dense nucleus. The electron “orbits” were at first imagined to be like the orbits of planets around the sun. Rutherford’s student, Niels Bohr, proposed that the electrons could orbit only in certain discrete orbits, but without radiating electromagnetic waves (e.g., light). This hypothesis was fruitful but contradicted what was and is known about electromagnetism. Louis de Broglie justified the discreteness by imagining electrom “waves”, of a fixed wavelength, forming wave patterns, such as those shown in the figures of Note 26, but of different radii to accommodate different numbers of wavelengths.

Thus, the radial component is an important part of the electron structure of the atom. This goes beyond this Week’s discussions of symmetry, so we won’t try to derive the functions. But it is helpful in visualizing the spherical harmonics we’ve just derived, so is worth the digression.

Finding the radial component requires some physics we don’t yet have. I will give the results for the “inverse square force” which describes both the gravitational attraction of planets by the sun and the electric (“Coulomb”) attraction of the orbital electrons by the nucleus.

Planets orbit the sun in ellipses, mostly all in the same plane, the “ecliptic”. (The confusing similarity of the two words is probably intentional.)

Electrons do not orbit in ellipses, as we will now see, despite the popular picture of the atom.

The radial component, $R_{n\ell}(r)$, depends on the radius r, on the angular momentum ℓ, which we have already seen in the spherical harmonics, and on the radial quantum number n, which is the modern form of de Broglie's quantized electron-wave levels.

Here, from [LL58, p.124], are the first few functions.

$$\begin{aligned}
R_{10}(r) &= 2e^{-r} \\
R_{20}(r) &= \frac{1}{\sqrt{2}}(1-\frac{r}{2})e^{-r/2} \\
R_{21}(r) &= \frac{1}{2\sqrt{6}}re^{-r/2} \\
R_{30}(r) &= \frac{2}{3\sqrt{3}}(1-\frac{2r}{3}+\frac{2r^2}{27})e^{-r/3} \\
R_{31}(r) &= \frac{8}{27\sqrt{6}}r(1-\frac{r}{6})e^{-r/3} \\
R_{32}(r) &= \frac{4}{81\sqrt{30}}r^2e^{-r/3}
\end{aligned}$$

Notice that $0 \le \ell < n$. From the experiment (Excursion for Note 26) with the playground roundabout, this makes some sense: the further out you are from the centre of rotation, the more angular momentum you can have.

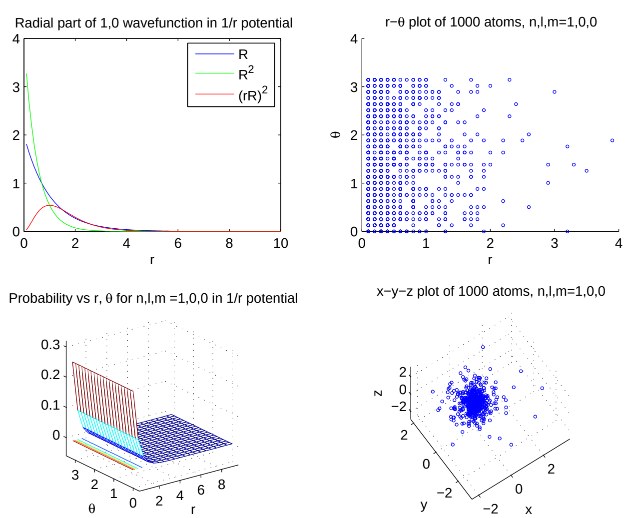
Notice that $R_{n\ell}(r)$ has *nodes*—values of r at which it goes to zero. You can see that there are $n-\ell-1$ of these nodes (apart from zeros at $r=0$):

none for R_{10}, R_{21}, R_{32}
one for R_{20} (at $r=2$), R_{31} (at $r=6$)
two for R_{30} (at $r=\frac{9}{2}(1\pm\frac{1}{\sqrt{2}})$)

Notice that the negative exponential ensures that $R_{n\ell}(r) \stackrel{r\to\infty}{\longrightarrow} 0$ so that the electrons do not normally escape from the atom.

Now let's draw and discuss some of these. The six plots to follow show a) $R_{n\ell}(r)$, b) $R^2_{n\ell}(r)$ combined with the spherical harmonic probability $Y_{\ell m}(\theta,\phi)Y^*_{\ell m}(\theta,\phi)$ (which does not depend on ϕ), and c) and d) 1000 atoms randomly generated under this probability and shown on the r-θ plane (c) and as a three-dimensional super-microscope view (d) of the "electron cloud" around a simple atom. In (c) and (d) each electron "shell"—separated from the other shells by $R_{n\ell}=0$ nodes—is shown in a different colour.

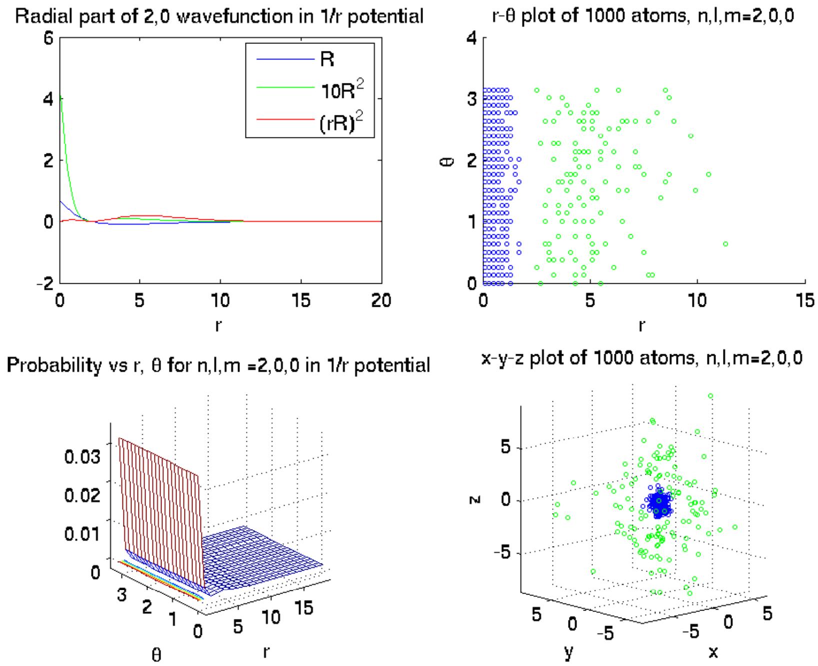
Here is $n,\ell,m = 1,0,0$



This has one shell (shown in blue) because there are zero nodal surfaces.

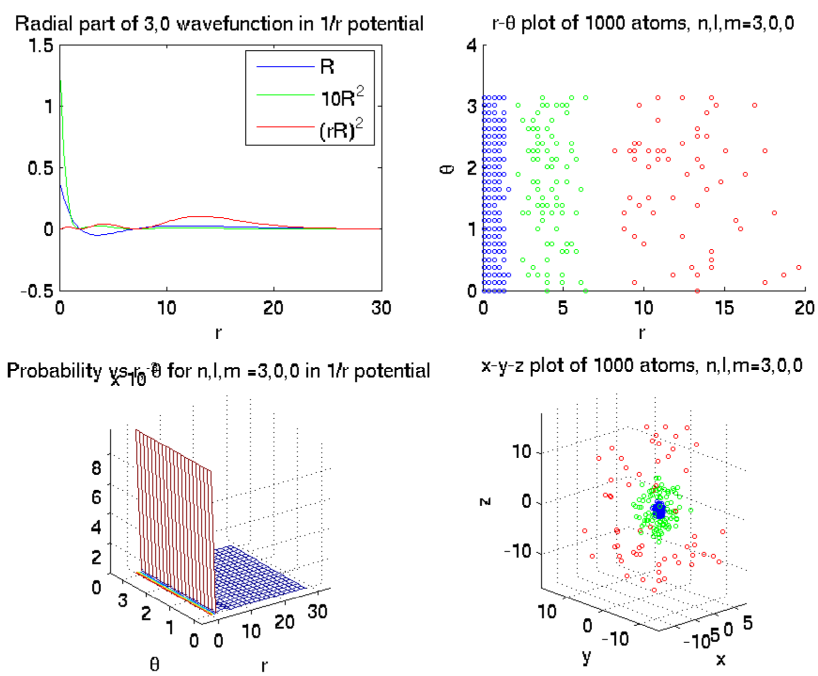
The electrons are distributed spherically, which is always the case when $\ell = 0$.

Here is $n, \ell, m = 2, 0, 0$



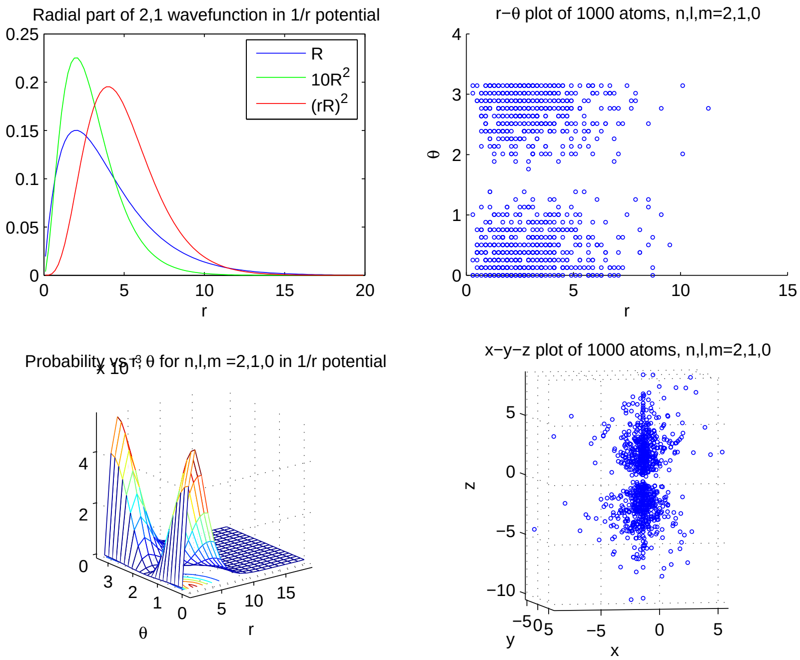
There are two shells (inner shell in blue, outer shell in green) because there is one nodal surface, at $r = 2$.

Here is $n, \ell, m = 3, 0, 0$



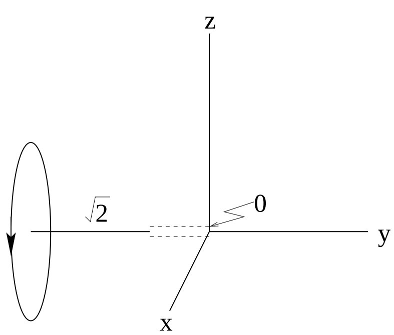
There are three shells (inner blue, middle green, outer red) and nodal surfaces at $r = 9(1-1/\sqrt{3})/2$ and $r = 9(1+1/\sqrt{3})/2$.

Here is $n, \ell, m = 2, 1, 0$

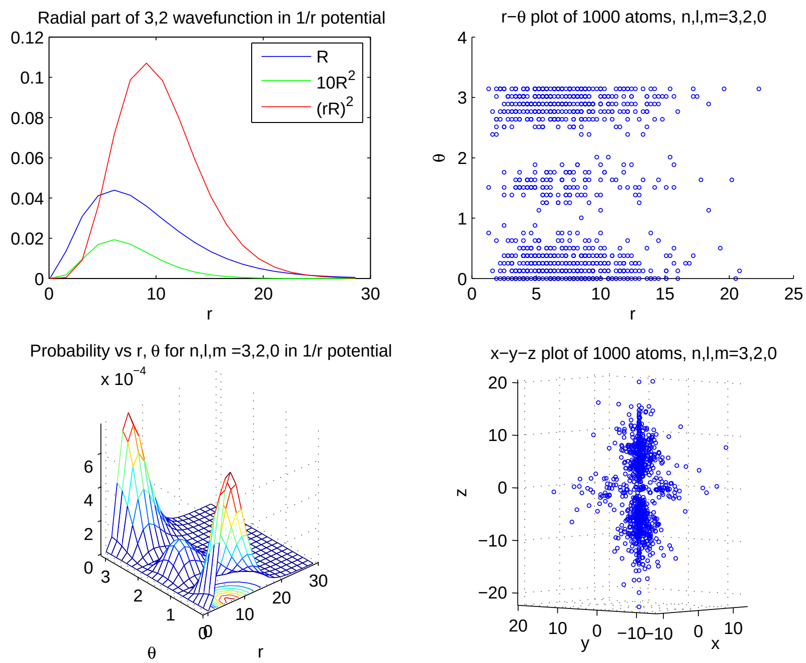


This has one shell but two clusters separated by the nodal plane (the equator) of $Y_{10}(\theta, \phi)$.

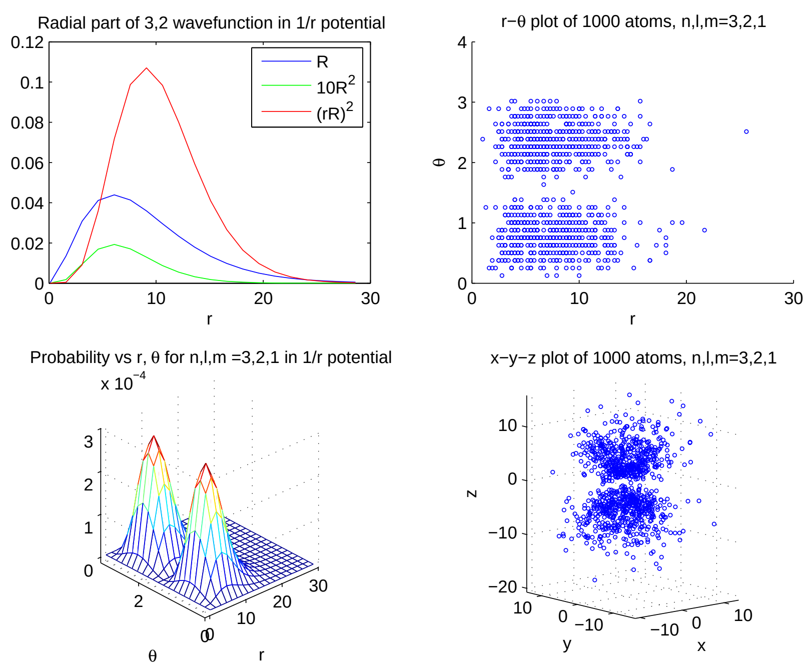
The angular momentum, $\ell = 1$, but the z-component, $m = 0$. We can visualize this (but the analogy is not perfect) as a top spinning perpendicularly to the z-axis. (Recall $\sqrt{\ell(\ell+1)} = \sqrt{2}$.)



Here is $n, \ell, m = 3, 2, 0$

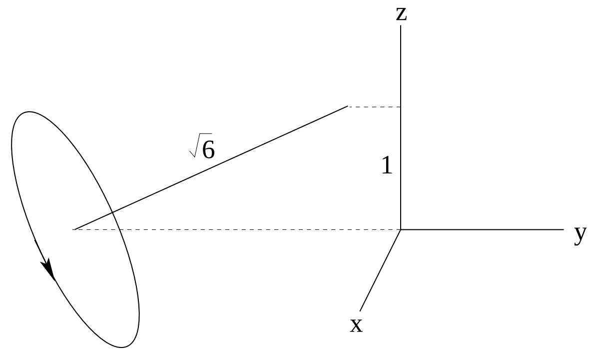


There is one shell but three clusters, separated by the two nodal planes of $Y_{20}(\theta, \phi)$.
Note that the spinning top picture for $n, \ell, m = 2, 1, 0$ only partly applies.
Here is $n, \ell, m = 3, 2, 1$



Compare this with $n, \ell, m = 2, 1, 0$ but note that the nodes are no longer exactly at the poles. We can see this from the plots other than the x-y-z plot. We would need to take cross−section of the x-y-z plot.

The spinning-top analogy is only suggestive here. (What angle must it be inclined at?) (Recall $\sqrt{\ell(\ell+1)} = \sqrt{6}$.)



The investigation of atoms that led to the understanding of angular momentum at atomic scales was spectroscopic. Electrons “falling” from one shell to a shell closer to the nucleus release energy in the form of light—sometimes ultraviolet or infrared—which can be analyzed by spectroscope. In hydrogen, for example, the so-called Balmer series is visible.

Angular momentum considerations give rise to fine structure in the spectra, which was characterized by the appearance of the lines of light in the spectroscope. Accordingly, the states of angular momentum got special names.

$\ell = 1$	s for “sharp”
$\ell = 2$	p for “principal”
$\ell = 3$	d for “diffuse”
$\ell = 4$	f for “fundamental”

and g, h, .. thereafter [Her44, pp.55ff.]

In summary, the possible electron states of an atom are

n	1	2				3								
ℓ	0	0	1			0	1			2				
m	0	0	1	0	−1	0	1	0	−1	2	1	0	−1	−2

and so on: n^2 states for the nth shell.

The different m states do not contribute to the spectra unless symmetry is broken, say by a magnetic or electric field in the z direction. These states are called ‘degenerate”.

The hydrogen atom is more completely described by SO(4,2), the special orthogonal group on four spacelike and two timelike dimensions. (The group of invariants under the Lorentz transformation is SO(3,1).)

32. SU(2) formal and informal,
33. SU(3).
34. Isospin and quarks
35. Symmetry and Conservation: Complementary Quantities
36. Symmetry and Conservation: Energy
37. Principle of Stationary Action.
38. Symmetry and Conservation: Noether’s Theorem
39. The Hamiltonian and Schrödinger’s Equation

40. Summary (These notes show the trees. Try to see the forest!)

Part I Discrete symmetries and molecules.

Notes 1–11. Symmetries of an equilateral triangle abstracted to groups. Invariant sets and subgroups. Traces and further matrix representations. Decomposing into irreducible representations and block-diagonalizing matrices.

Notes 12–17. Finding fundamental vibration modes of molecules from their symmetries: greenhouse gases CO_2 and H_2O.

Notes 16–18. Symmetries of the platonic solids: tetrahedron, octahedron/cube, dodecahedron/icosahedron.

Part II Infinite symmetries and crystals.

Notes 19–25. Translation symmetries and crystals in one and two dimensions: crystallography and waves.

Part III Continuous symmetries and the atom.

Notes 26–29. Rotational symmetry in two and three dimensions. Commutator algebra and representations of the spherical group.

Notes 30, 31. Spherical harmonics and atomic physics.

Part IV Abstract symmetries and lots of physics

Notes 32–34. From SU(2) (the atom) to SU(3) (the quarks). Isospin and hypercharge.

Notes 35–39. Symmetry and conservation laws. Complementary quantities, energy, Lagrangian, principle of stationary action, Noether's theorem, Hamiltonian, Schrödinger's equation and the quantum harmonic oscillator.

II. The Excursions

You've seen lots of ideas. Now *do* something with them!

1. Look ahead to Week 10 (Excursion "More than one kind of infinity") and Week 12 (Excursion "Continuity") to refine the discussion in Note 26 about continuous groups.

2. a) Write the programs to check the series for $\cos x$ and $\sin x$.
   b) Remembering from Week 8 (Excursion "Slopes of cos and sin") that
   $$\text{slope} \cos x = -\sin x \qquad \text{and} \qquad \text{slope} \sin x = \cos x$$
   check these with the slopes of the two series in Note 26.
   c) What are slope slope cos, slope slope slope cos, etc. and the same for sin? How do these compare with the same repeated slopes of the two series?
   d) If the series for cos and sin did not alternate signs, could they still be periodic ($\cos 2\pi = \cos 0$)?
   e) What are the slopes of e^{ix} and its series?
   f) What is the series for e^x? What is its slope?

3. a) Diagonalize
   $$R = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$
   by QRQ^{-1} where
   $$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$$
   b) Find QXQ^{-1} for
   $$X = -i \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$$
   c) Should the same process diagonalize both X and e^{iX}, given the definition of e^{iX}?

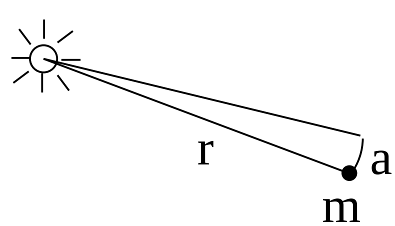
4. Show the following concerning the circular waves plotted in Note 26.
   a) If m is not an integer, the wave will interfere destructively with itself everywhere, giving zero amplitude.
   b) If the symmetry were not circular but had n discrete points, higher frequencies (values of m) are not distinguishable from lower frequencies, above a certain frequency, because the n nodes move the same way for the higher frequencies as for the lower. What is the highest value of m that gives a distinct vibration?

5. To experience the conservation of angular momentum you can try the experiment with a bicycle wheel and swivel chair, if you have each of these, especially a bicycle wheel you can spin while holding the axle, one hand at each end, without getting too greasy.
   Sit on the chair, holding the wheel axle horizontally and have a friend spin the wheel. Sitting securely but with your feet off the ground, rotate the axle until it is vertical. What happens? Why?
   If you don't have the above research equipment, here is a possibly dangerous alternative if you can find one of those playground rotating platforms, built quartered with four steel handrails to hang on to.
   With one or three friends of your own weight and size distributed on the platform, one opposite you and, if there, the other two opposite each other symmetrically on each side of you, all leaning as far away from the centre as they can while holding on tightly with two hands, spin the platform up by *walking* around it, then get on yourself, leaning outward and holding on tightly with both hands.
   Now, all on the slowly rotating platform, tell everyone to pull themselves in toward the centre¿ What happens? Why?
   Do not get off without everybody leaning away from the centre again, so the platform is going slowly enough for one person to step off and bring it to a halt.
   Be careful with the experiment! Do not try it with friends much smaller or much bigger than you.

6. (Ron Niemi.) Why does a tightrope artist hold a long pole out from side to side, or somebody doing a lesser act of balancing hold out heir arms?

7. The above experiments with angular momentum show that it involves direction and motion. The discussion of Note 26 involves only a standing wave. How can a standing wave have direction or motion?
   Look up "standing wave" and "surf" to see. For instance, try the YouTube video "Waimea River Standing Wave Surf Session".
   Find a standing wave near you (e.g., the Habitat wave in the St. Laurence River or the Chambly wave in the Richelieu River, both near Montreal) and surf on it to appreciate better both the feel of momentum and my paraphrase of T. H. Nelson's remark: if computers are the wave of the future, graphics is the surfboard.

8. **Kepler II.** Show that a body of mass m orbiting a central body with constant angular momentum, J, sweeps out equal areas in equal times at rate $J/(2m)$ area units per time unit (or $2m/J$ time units per area unit) using the pre-quantum expression for angular momentum $J = mvr$ where v is the velocity of the orbiting body and r its distance from the central body.



(Suppose the orbiting body travels along an arc of length a in a time so short that a is very small and the area swept out can be considered a triangle of base r and height a.)

9. Show that the J' matrices have the commutators shown in Note 27.

10. Show that $(\vec{\alpha}.\vec{J}')\begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and so $(\vec{\alpha}.\vec{J}')^3$ is as shown in Note 27.

11. In Week 6 Note 3 we gave a 2D representation of the spherical group using Euler angles. Let's now give it in terms of angle ϕ about an arbitrary axis, as we do in Note 27 for the 3D representation.
a) Show that the half-Pauli matrices

$$J_x = \frac{1}{2}\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad J_y = \frac{1}{2}\begin{pmatrix} & -i \\ i & \end{pmatrix}, \quad J_z = \frac{1}{2}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

have the same commutator relationships as the Js of Note 27.
b) Show that

$$e^{i\vec{\phi}.\vec{J}} = I\cos\frac{\phi}{2} + \begin{pmatrix} ir & ip+q \\ ip-q & -ir \end{pmatrix}\sin\frac{\phi}{2}$$

where

$$\vec{\phi}\cdot\vec{J} = \phi p J_x + \phi q J_y + \phi r J_z$$

c) Check that this result agrees with Week 6 Note 3 for rotations through ϕ about each of the x, y and z axes (Week 6 Note 6).
d) What is the character table for this 2D representation? (All rotations by a fixed angle ϕ form an invariant class. Why?)
e) What is the character table for the 3D representation of Note 27?

12. Show that $\ln(e^x) = x$ and $e^{\ln(x)} = x$ using the series for $e^x = e^{i(x/i)} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}..$ and the series for $\ln(x)$ given in Note 28.

13. Show that $[X, X^\dagger] = 0$ for hermitian X (i.e., $X = Y + iZ,\ Y^T = Y,\ Z^T = -Z$).

14. Show that $[J'_x, [J'_y, J'_z]] + [J'_y, [J'_z, J'_x]] + [J'_z, [J'_x, J'_y]] = 0$ for the generators of Note 27.

15. What I've called a commutator algebra in Note 28, to emphasize the idea from which it was abstracted, is properly called a Lie algebra, invented by Sophus Lie. Continuous groups with generators whose commutators obey the axioms of Lie algebra are called Lie groups. Look up Sophus Lie, 1842–99.

16. Show that the J matrices in Note 29 have the same commutator relationships as the J' matrices.

17. a) Show that $(2\ell_1 + 1)(2\ell_2 + 1) = (2(\ell_1 + \ell_2) + 1) + \ .. + \ (2 \mid \ell_1 - \ell_2 \mid +1)$ in $2\min(\ell_1, \ell_2) + 1$ steps.
b) Complete the decompositions in the (ℓ_1, M_1) by (ℓ_2, M_2) table in Note 29.

18. a) Show that each rotation angle α specifies an invariant class of elements in the spherical group, and so that the characters of the 3D representation are $1 + 2\cos\alpha$.
b) What are the group elements and hence the characters in the 2D representation? 4D?
c) What are the invariant subgroups?

19. For what values of ℓ does $\sqrt{\ell(\ell+1)}$ equal ℓ to within one part in a million? (Note 29)
For what values of ℓ do two successive values, ℓ and $\ell + 1$, differ by less than one part in a million?
What are the smallest values of n (Note 31) that allow these values of ℓ?

20. Show that the 4D generators from Note 29,

$$(p,q,r)\cdot\vec{J} = \begin{pmatrix} \frac{3}{2}r & \frac{\sqrt{3}}{2}(p-iq) & & \\ \frac{\sqrt{3}}{2}(p+iq) & \frac{1}{2}r & p-iq & \\ & p+iq & \frac{1}{2}r & \frac{\sqrt{3}}{2}(p-iq) \\ & & \frac{\sqrt{3}}{2}(p+iq) & \frac{3}{2}r \end{pmatrix}$$

has the same commutator relationships as the J's of Note 27.

21. Confirm the rules in the "Slope Rules" table in Note 30.
a) Look up the binomial coefficients in Week_ii Notes 6 and 7 and show the power rule using the definition

$$\text{slope}_x f(x) \approx \frac{\Delta f}{\Delta x} = \frac{f(x+\Delta x)-f(x)}{\Delta x}$$

for small shifts, Δx, in the value of x.
b) Do Week 8 Excursion "Slopes of cos and sin" to show the trig. rules.
c) Check this demonstration of the product rule:

$$\begin{aligned} \Delta(fg) &= f(x+\Delta x)g(x+\Delta x) - f(x)g(x) \\ &= f(x+\Delta x)g(x+\Delta x) - f(x)g(x+\Delta x) + f(x)g(x+\Delta x) - f(x)g(x) \\ &= (\Delta f)g(x+\Delta x) + f(x)D\delta g) \\ &\approx (\Delta f)g(x) + f(x)(\Delta g) \end{aligned}$$

d) Check this demonstration of the first chain rule:

$$\begin{aligned} \frac{\Delta f(g)}{\Delta x} &= \frac{f(g(x+\Delta x)-f(g(x)}{\Delta x} \\ &= \frac{f(g+\Delta g)-f(g)}{\Delta x} \\ &= \frac{f(g+\Delta g)-f(g)}{(g+\Delta g)-g}\frac{(g+\Delta g)-g}{\Delta x} \\ &= \frac{f(g+\Delta g)-f(g)}{\Delta g}\frac{(g+\Delta g)-g}{\Delta x} \\ &\approx \text{slope}_g f\ \text{slope}_x g \end{aligned}$$

Combine this argument with that of (c) to show the second chain rule.
e) $f^{-1}(x)$ is defined to be the inverse function to $f(x)$:

$$x = f^{-1}(f(x)) = f(f^{-1}(x))$$

(Not every function has an inverse everywhere.)
Use the chain rule to show that

$$1 = \text{slope}_y f^{-1}(y)\ \text{slope}_x f(x)$$

and hence show the inverse rule.
f) Use the inverse rule to show the two inverse trig rules. Explain the tricky part of this.

22. Confirm $[J_x, J_y] = x\ \text{slope}_y - y\ \text{slope}_x$ if

$$\begin{aligned} J_x &= -i(y\ \text{slope}_z - z\ \text{slope}_y) \\ J_y &= -i(z\ \text{slope}_x - x\ \text{slope}_z) \end{aligned}$$

Remember how some of the terms must work out, such as

$$y\ \text{slope}_z z\ \text{slope}_x = y\ \text{slope}_x + yz\ \text{slope}_z\ \text{slope}_y$$

23. In calculating $Y_{\ell m}$ in Note 30 we show only integer ℓ and m. The theory also gives half-integer ℓ and m. Derive these. They are left out of Note 30 because they have almost no effect on atoms (Note 31).

24. Find the ℓ nodal lines in the spherical harmonics figures in Note 30.

25. Run and understand the MATLAB program (see `spherHarm.m` in MATLABpak08cIII) that generated the spherical harmonics figures in Note 30. Explore these plots in 3D and try some of the other spherical harmonics. Look up still more spherical harmonics and extend the program.

26. How will Y_{00} $(\ell, m = 0, 0)$ appear if drawn in the same way as the spherical harmonics of Note 30?

27. Why is the probability, $Y_{\ell m}Y^*_{\ell m}$, for the spherical harmonics in Note 30 always independent of ϕ?

28. Why is the imaginary part always 0 for the spherical harmonic $Y_{\ell\,0}$ $(m = 0)$?

29. a) Show directly that $J_-Y_{\ell,-\ell}(\theta, \phi) = 0$ is satisfied by $\sin^{-\ell}\theta e^{-i\ell\phi}$ (times some normalizing constant depending only on ℓ).
b) Hence show that ℓ, instead of taking on any value, which would formally satisfy the equation, may only be an integer or a half-integer.

30. Compare the atomic plots in Note 31 with the graphs and renditions in [Her44, pp.40,43,44] (which I cited in Week 6 and which you are now in a position to understand a lot better).

31. Run and understand the MATLAB program (see `atom.m` in MATLABpak08cIII) that generated the figures of atomic structure in Note 31. Explore the 3D plots from various angles. Look up more $R_{\ell,m}(r)$ functions and extend the program.
How would you modify the program to produce a cross-section of the "electron cloud"?

32. How will $n, \ell, m = 2, 1, 1$ appear if drawn the same way as the atomic figures of Note 31? How would its spinning top analogy look?
Show that $n, \ell, m = 3, 1, 1$ will be similar, but with two shells.

33. Why will $n, \ell, m = 3, 1, 0$ look like $n, \ell, m = 2, 1, 0$ but with two shells?

34. **Bohr atom.** Find the "Bohr radius" $a_0 = \hbar^2/(m_e k_H) = 52.9$ picometers, and the "Rydberg constant" $R = k_H/(2a_0) = 13.6$ eV, using

	constant	"dimensions"	name
$\hbar$	$=6.582_{10}{-16}$ eV-sec	ET	Planck's constant
m_e	$=0.511_{10}6$ eV	E	electron mass
k_H	$=8.68_{10}{-10}$ eV-m	EL	hydrogen-atom potential energy factor

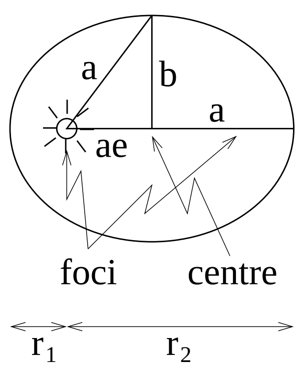
(For "dimensions" see Week 7a Notes 7, 8: E energy, T time, L length, M mass. An electron-Volt, eV, is the energy given an electron by a 1-Volt potential difference, and is a suitable unit for atom-scale processes.)
These numbers have the following significance for the radii, r_n of the electron shells in the atom and for the wavenumbers ν_{nm} of light emitted by an electron falling from shell n to shell m $(n > m)$.
Given the momentum of the electron, $p = h/\lambda_e$, and that the de Broglie wavelength λ_e of the electron must fit an integer number of times in its orbit at radius r, $n\lambda_e = 2\pi r$, we see $p = n\hbar/r$ where $\hbar = h/2\pi$.

Now we invoke both the conservation of pre-quantum angular momentum, $p_1 r_1 = J = p_2 r_2$, and of pre-quantum energy

$$\frac{p_1^2}{2m_e} - \frac{k_H}{r_1} = E = \frac{p_2^2}{2m_e} - \frac{k_H}{r_2}$$

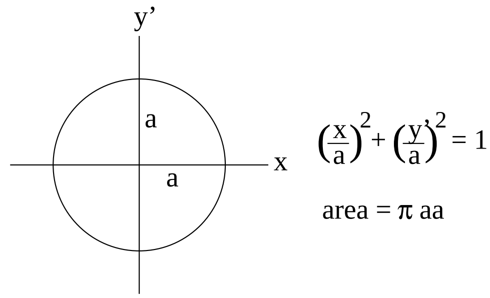
for two different points and momenta on the electron "orbit", r_1, p_1 and r_2, p_2.
Using $\frac{1}{r_1^2} - \frac{1}{r_2^2} = (\frac{1}{r_1} + \frac{1}{r_2})(\frac{1}{r_1} - \frac{1}{r_2})$ show that $J^2(\frac{1}{r_1} + \frac{1}{r_2}) = 2m_e k_H$
Making these the same point again ($r_1 = r_2,\ \ p_1 = p_2$), $J^2/r = m_e k_H$,
i.e., $p^2/(2m_e) = k_H/(2r)$.
Do two things with this. First $n^2\hbar^2/m_e = k_H r$ so $r = \hbar^2 n^2/(m_e k_H) = a_0 n^2$.
Second $E = k_H/(2r) - k_H/r = -k_H/(2r) = -k_H/(2a_0 n^2) = -R/n^2$.
So the "orbital radii" are proportional to n^2 and the smallest radius of the hydrogen atom is 0.0529 nm.
So an electron "falling" from shell n to shell m gives off energy $E_{nm} = R(1/m^2 - 1/n^2)$ corresponding to photon wavenumber (per cycle)
$\nu_{nm} = f_{nm}/c = E_{nm}/(hc) = 10.97(1/m^2 - 1/n^2)$ million waves per meter.
What are the photon wavelengths for the first few terms of the Balmer series, in which $m = 2$, in hydrogen spectroscopy?
The factor of 2 that appears in the relationships among total energy, $E = -k_H/(2r)$, kinetic energy, $\mathcal{T} = p^2/(2m_e) = k_H/(2r)$, and potential energy, $\mathcal{V} = -k_H/r$, is not a coincidence but is generally true for stable systems bound by potential energy, with 2 always appearing for inverse-square forces. Look up the "virial theorem".

35. **Kepler III** a) Use the previous Excursion and the Excursion "Kepler II" to show that the time required for an entire orbit (its "period") in a circle of radius r is T where $T^2 = 4\pi^2 r^3/k_S$ where k_H for the hydrogen atom has been replaced by k_S for the solar system, $k_S = 23.8$Gsolday. (A solday is the amount of energy put out by our sun in one day, and is a suitable unit for cosmic-scale processes.)
b) If the circle were replaced by an ellipse of semi-major axis a, show that the above expression holds with r replaced by a, no matter what the semi-minor axis b is.

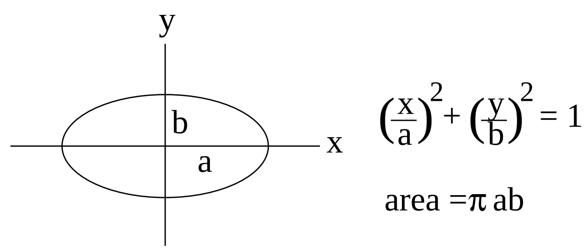


(Use r_1 as the shortest and r_2 as the longest distances from the sun (nucleus) and show that $r_1 r_2 = b^2$ and $r_1 + r_2 = 2a$.
Using the definition of an ellipse as the curve made by a pencil constrained to keep taut a constant-length string with ends attached to each focus, show that the two as in the figure are indeed the same.)

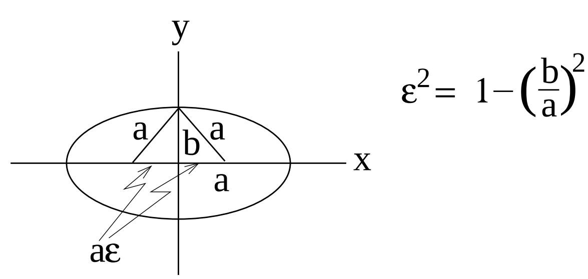
36. **Conic sections.** 1) Circle:



2) We almost never see circles straight-on but usually from an angle: what we are really seeing is an *ellipse*, a squashed circle. Try $y = by'/a$:



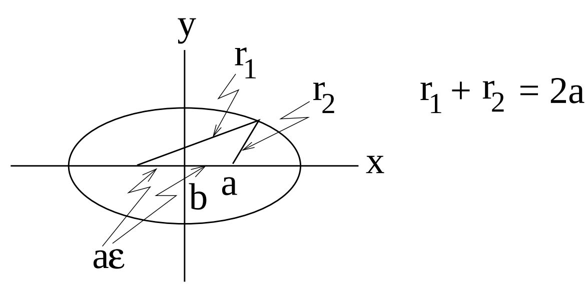
(So, for future reference, $y^2(a/b)^2 = a^2 - x^2$.)
3) Now introduce a quantity ϵ (written slightly differently in the diagrams for this Excursion) which measures how far away from the circle we have gone. It could be b/a, but that is 1 for the circle ($b = a$) and it might be better to make a measure which is 0 for the circle.
What turns out to be handy, although not immediately obvious, is to define $\epsilon^2 = 1 - (b/a)^2$. To justify this, let's find the two points on the x-axis a distance $a\epsilon$ from the centre, and call them the *foci* (plural of *focus*) of the ellipse:



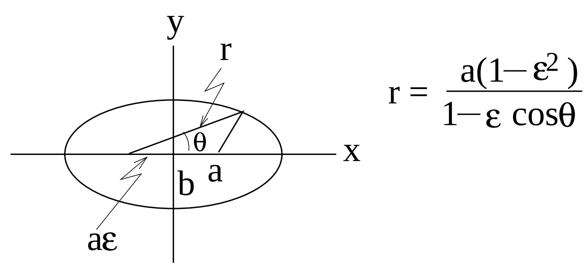
Then the definition of ϵ tells us that the hypotenuses of the triangles shown are a. (And $y^2 = (b/a)^2(a^2 - x^2) = (1 - \epsilon^2)(a^2 - x^2)$.)
4) Now consider any point on the ellipse, distances r_1 and r_2 from the two foci. (These are not the same r_1 and r_2 as the previous Excursion.) A little algebra:

$$\begin{aligned} r_1^2 &= (x + a\epsilon)^2 + y^2 = (a + \epsilon x)^2 \\ r_1^2 &= (x - a\epsilon)^2 + y^2 = (a - \epsilon x)^2 \end{aligned}$$

So $r_1 + r_2 = 2a$ for any point on the ellipse.

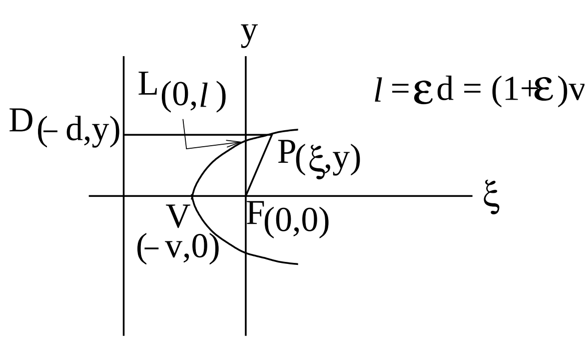


This gives the "string method" of describing an ellipse: put two pins at the foci, tie a loop of string loosely around them, put a pencil in so that the string is taut between pencil and both pins, and draw the curve you get by always keeping the loop tight.
5) We can get the polar equation for the ellipse, in terms of distance r from one focus, and angle θ to the x-axis at that focus.



$$\begin{aligned} x &= r\cos\theta - a\epsilon \\ r &= a + \epsilon x = a + \epsilon(r\cos\theta - a\epsilon) \\ r &= \frac{a(1-\epsilon^2)}{1-\epsilon\cos\theta} \end{aligned}$$

Check that this gives the right answers for $\theta = 0,\ \cos^{-1}\epsilon$ and π. (When $\cos\theta = \epsilon$ what are the values of x and r?)
Plot the ellipse using polar coordinates and suitable values for a and ϵ $(0 \le \epsilon < 1)$.
What happens if $-1 < \epsilon \le 0$?
What happens if $\epsilon > 1$ or $-1 < \epsilon$?
6) What happens if $\epsilon = 1$?
Squaring both sides of $r = a + \epsilon x$ gives $(x + a\epsilon)^2 + y^2 = (a + \epsilon x)^2$ or $x^2(1-\epsilon^2) + y^2 = a^2$, which is not much help when $\epsilon = 1$.
So we take a new tack and redefine the curve as all points $P = (x, y)$ such that FP $= \epsilon$DP



This give the relationship in the figure, $\ell = \epsilon d = (1+\epsilon)v$.
(The point F = (0,0) is the *focus*, the line D $= (-d, y)$ is the *directrix*, the point V $= (-v, 0)$

is the *vertex* and the length 2ℓ is the *latus rectum.*)
For an ellipse $v = a(1-\epsilon)$ so $\ell = a(1-\epsilon^2) = b^2/a$, which is the numerator in the polar equation.
So we rewrite that equation as

$$r = \frac{\ell}{1 - \epsilon\cos\theta}$$

In general, FP $= \epsilon$DP gives, when squared, $\xi^2 + y^2 = \epsilon^2(\xi + d)^2$ or $\xi^2(1-\epsilon^2) - 2\epsilon^2 d\xi + y^2 = \epsilon^2 d^2$.
(For an ellipse, using $d = a(1-\epsilon^2)/\epsilon$ and shifting $\xi \rightarrow x = \xi - a\epsilon$,

$$(x + a\epsilon)^2(1-\epsilon^2) - 2a\epsilon(1-\epsilon^2)(x + a\epsilon) + y^2 = a^2(1-\epsilon^2)^2$$

or $x^2 + (ay/b)^2 = a^2$.)
Now try $\epsilon = 1$: $-2d\xi + y^2 = d^2$ or $\xi = y^2/(2d) - d/2 = y^2/(4v) - v$
This is a parabola with focal length v and semi-latus rectum $\ell = d = 2v$, twice the focal length.

37. **Kepler I.** ($1/r$ potentials give motion in conic sections.)
In the Excursion on conic sections we found $r = \ell/(1 - \epsilon\cos\theta)$ for a conic section with semi-latus rectum ℓ and eccentricity ϵ.
a) For reasons that will be clear in (e), show that

$$\text{slope}_r\theta = \frac{\ell/r}{\pm\sqrt{(\epsilon^2 - 1)r^2 + 2\ell r - \ell^2}}$$

Hint. Since $\cos\theta = (r-\ell)/(r\epsilon)$, equate $\text{slope}_r \cos\theta = -\sin\theta\ \text{slope}_r\theta$ with $\text{slope}_r(r-\ell)/(r\epsilon)$
b) Since angular momentum is conserved by central forces such as the Bohr atom or the sun-planet attraction in the solar system, the Kepler II Excursion says that equal areas are swept out in equal times at rate $J/(2m)$. This restricts planetary motion to a plane interval which we can call $e_{12}J/(2m)$ in the interval algebra (Week 7c Notes 6–11).
In pre-quantum physics we may use velocities instead of momenta, $p = mv$ or $J = mvr$, and velocity is the slope, with respect to time, of position, $v = \text{slope}_t x$.
(In quantum physics, velocity cannot be defined because we cannot know both the position and momentum simultaneously: "successive positions" is an undefined concept.)
In the two dimensions of the planetary orbit we write the angular momentum in the interval algebra, $J = m\ \mathbf{cmpt}(2,\, q\dot{q})$ where position

$$q = xe_1 + ye_2 = r\cos\theta e_1 + r\sin\theta e_2$$

and velocity

$$\begin{aligned}\dot{q} &= \text{slope}_t q \\ &= (\dot{r}\cos\theta - \dot{\theta}r\sin\theta)e_1 + (\dot{r}\sin\theta + \dot{\theta}r\cos\theta)e_2\end{aligned}$$

(check this) where the dotted variables are slopes with respect to time of the corresponding undotted variables.
Show that $J = m\dot{\theta}r^2 e_{12}$
Since there is only the e_{12} component, we will drop the e_{12} for the rest of this Excursion and write

$$J = m\dot{\theta}r^2 \tag{1}$$

Why is this the same as $J = mvr$ which we used before?
c) Kinetic energy is $\mathcal{T} = m\dot{q}\dot{q}/2 = m(\dot{r}^2 + \dot{\theta}^2 r^2)/2$. Show that $\dot{q}\dot{q} = \dot{r}^2 + \dot{\theta}^2 r^2$.
d) Potential energy in the solar system is $\mathcal{V} = -k_S/r$, where $k_S = 23.8$Gsolday (see the Kepler

III Excursion).
The total energy, $E = \mathcal{V} + \mathcal{T}$, can be written in terms of an "effective potential", $J^2/(2mr^2) + \mathcal{V}$. Show this by replacing $\dot{\theta}$ by $J/(mr^2)$ (from equation 1) in the equation for $\mathcal{T}$.
(The extra term gives rise to the "centrifugal force", $-\text{slope}_r J^2/(2mr^2) = J^2/(mr^3) = mv^2/r$. But forces are beyond what we understand so far.)
Thus

$$E = \frac{m}{2}(\dot{r}^2 + \frac{J^2}{m^2r^2}) - \frac{k}{r} \tag{2}$$

e) Show that equations 1 and 2 give

$$\begin{aligned}\frac{\dot{\theta}}{\dot{r}} &= \frac{J/(mr)}{\pm\sqrt{2E/m + 2k_S/(mr) - J^2/(m^2r^2)}} \\ &= \frac{J^2/(k_S mr)}{\pm\sqrt{2EJ^2r^2/(k_S^2 m) + 2J^2 r/(k_S m) - (J^2/(k_S m))^2}}\end{aligned}$$

and so, comparing with (a) and because $\text{slope}_t\theta = \text{slope}_f\theta \; \text{slope}_t r$ by the chain rule (Note 30)

$$\begin{aligned}\ell &= \frac{J^2}{km} \\ \epsilon &= \pm\sqrt{1 + \frac{2EJ^2}{k_S^2 m}}\end{aligned}$$

This shows that a $1/r$ central potential gives rise to motion in conic sections—ellipses if the total energy is negative, parabolas if it is zero, and hyperbolas if it is positive.
(Newton told Halley: "why, I have calculated it".)
At least, the differential equations relating r and θ are the same. For a more thorough discussion along these lines (using Gibbs' vector notation) see [Bae03].
In order to make his *Philosophi Naturalis Principia Mathematica* acceptable to a readership without calculus, Newton gave a geometrical demonstration of Kepler I. Look up Feynman's reconstruction of this geometrical argument.

38. What difference in behaviour in the (quantum) atom marks the distinction between negative and positive energies that gives rise to elliptic versus hyperbolic orbits in the solar system?

39. Any part of the Preliminary Notes that needs working through.

References

[Bae03] John Baez. Mysteries of the gravitational 2-body problem. URL math.ucr.edu/home/baez/gravitational.html Accessed 09/2/17 A linked homework pdf treats the Kepler problem., May 3 2003.

[Her44] Gerhard Hertzberg. *Atomic Spectra and Atomic Structure*. Dover Publications, New York, 1944. Prentice-Hall Inc., 1937; translated by J. W. T. Spinks.

[LL58] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics: Non-relativistic Theory*. Pergamon Press Ltd, Oxford, 1958. Vol 3 of Course of Theoretical Physics; Translated by J. B. Sykes and J. S. Bell.