Excursions in Computing Science:
Book 11d. Forces and Invariants
Part IV. Quantum Field Theory: Matrix Quantum Mechanics

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

August 1, 2021

I. Prefatory Notes

18. Introduction to Quantum Fields. We have been following Schrödinger in describing particles as wavefunctions, whose values in time and space are amplitudes, the squares of which are the

probabilities that a particle is present at the given time and place. Schrödinger says nothing about how the particle came into existence in the first place or what happens to it in the end.

At sufficiently high energies, special relativity comes into play and particles can be created or destroyed. This is because $E = mc^2$ says that energy and mass are equivalent and so if there is sufficient energy, E, a particle of mass m can come into being. Particle physics, or high-energy physics, is largely about the appearance and disappearance of particles.

Here are some bubble-chamber tracks of electrons and positrons appearing at the point of disappearance of photons.

Since the photon is not charged it leaves no track in the bubble chamber. But at the vertex of each V in the image, a photon has disappeared and an electron-positron pair appears. They are charged and so make tracks. Their charges are opposite and since a magnetic field has been applied to the bubble chamber, these tracks curve in opposite directions. Since they have the same masses and velocities, and the same charges but with opposite signs, these curvatures are equal but opposite (wherever the magnetic field is uniform).

Evidently we are going to need some operators which make particles appear or disappear. Let's call them U for "uppear" and D for "disappear". (My reason for distorting the terminology will become clear in the mathematics of the next section.)

Historically these have been called creation and annihilation operators, respectively, or emission and absorption operators, respectively. (You can see that using A instead of U might be confusing.)

We can begin to explore the properties of the operators U and D. They will need subscripts to identify the particles they apply to, and we'll see that they will also need symbols to distinguish which *types* of particle they apply to—e.g., photons are a different type from electrons or positrons.

Feynman [FLS64, pp.III-4-1 f.] shows that, because quantum physics deals in amplitudes rather than probabilities, all particles must come in one of two types: those whose amplitude signs change when two identical particles are swapped, and those whose signs don't. (See also Week 5, Notes 8 and 9.) Or, we could reflect a single particle: the amplitude on reflection might change by a phase factor $e^{i\theta}$, but the square of this phase factor must be 1 because two reflections give the identity. Thus $e^{i\theta} = \pm 1$ for which there are two possibilities, hence two basic types of particle.

If we create two identical particles, we could use either U_1U_2 or U_2U_1, applied to something, to generate amplitudes. (I've distinguished the particles by subscripting the Uppear operator, although physically they are not distinguishable.) Then for the $+1$ type of particle

$$U_2U_1 = U_1U_2$$

but for the -1 type of particle

$$U_2 U_1 = -U_1 U_2$$

We can express this difference by using *commutators*

$$[U_{(-)1}, U_{(-)2}]_- \stackrel{\text{def}}{=} U_{(-)1}U_{(-)2} - U_{(-)2}U_{(-)1}$$

and *anticommutators*

$$[U_{(+)1}, U_{(+)2}]_+ \stackrel{\text{def}}{=} U_{(+)1}U_{(+)2} + U_{(+)2}U_{(+)1}$$

And I've labelled the U operators with the corresponding signs.

This same discussion applies to distinguish $D_{(-)}$ from $D_{(+)}$.

All particles described by the commuting operators are called Bose-Einstein particles or *bosons*.

All particles described by the anticommuting operators are called Fermi-Dirac particles or *fermions*.

Commuting operators such as U_- could possibly be represented by ordinary numbers—although we're about to meet other aspects which do not commute—but the anticommuting operators are new. The mathematical objects we know that come closest to this are *matrices*. although we have yet to explore anticommuting matrices.

Commutators first: what arises is the question, how do *different* operators interact, e.g., U with D? We start by looking at H and A, and at X and P.

Before Schrödinger formulated quantum mechanics as a slope equation

$$i\hbar\partial_t u = -\frac{\hbar^2}{2m}\partial_x^2 u + Vu = Hu$$

(Notes 10 and 11 above), Heisenberg had already worked out a form of *matrix mechanics*. So we should look at that. Schrödinger puts the time dependence into the state vector (or function) u, which we'll now call u_S. Heisenberg puts the time dependence into the operators, leaving what we'll call u_H independent of time. Formally, then,

$$u_S = e^{-iHt/\hbar}u_H$$

as we saw in Note 11: since $\partial_t u_H = 0$, this solves the Schrödinger equation

$$i\hbar\partial_t u_S = Hu_S$$

So the time dependence is all contained in the operator in the Heisenberg picture. Thinking of an operator A as a matrix, the transformation corresponding to that for the vector u, is

$$A_S = e^{-iHt/\hbar}A_H e^{iHt/\hbar}$$

or

$$A_H = e^{iHt/\hbar}A_S e^{-iHt/\hbar}$$

The time dependence of a Heisenberg operator, given $\partial_t A_S = 0$, is

$$i\hbar\partial_t A_H = A_H H - HA_H = [A_H, H]_-$$

So $\partial_t A_H = (i/\hbar)[H, A_H]_-$. This is the Heisenberg equivalent, for operator A, to the Schrödinger equation for wave function u.

Commutators are beginning to seem important. Here's another. In Note 10 we used the equivalence $p = -i\hbar\partial_x$ (one dimension is enough for this discussion). So $[x, p]_-$ applied to some u is

$$\begin{aligned}
[x,p]_- u &= [x, -i\hbar\partial_x]_- u \\
&= -xi\hbar\partial_x u + i\hbar\partial_x xu \\
&= -xi\hbar\partial_x u + i\hbar u + xi\hbar\partial_x u \\
&= i\hbar u
\end{aligned}$$

or, in short

$$[X, P]_- = i\hbar$$

I've capitalized P and X because they are both operators.

So what do they look like as matrices?

Here's a trick. Let's replace P and X by two linear combinations from which we can easily get P and X back. And let's replace the constant $\hbar$ by 2 just to make the math really simple: we can always put $\hbar/2$ back at the end.

$$\begin{aligned} D &\stackrel{\text{def}}{=} \frac{1}{2}(X + iP) \\ U &\stackrel{\text{def}}{=} \frac{1}{2}(X - iP) \end{aligned}$$

So

$$\begin{aligned} [D, U]_- &= \frac{1}{4}((X + iP)(X - iP) - (X - iP)(X + iP)) \\ &= \frac{i}{4}(PX - XP + PX - XP) \\ &= \frac{-i}{2}[X, P]_- \\ &= I \end{aligned}$$

Why am I doing this? Bear with me one more step. If we suppose D and U are "ladder" matrices, written respectively as a single superdiagonal and a single subdiagonal

$$DU = \begin{pmatrix} a & & & & \\ & b & & & \\ & & c & & \\ & & & . & \\ & & & & . \end{pmatrix} \begin{pmatrix} A & & & & \\ & B & & & \\ & & C & & \\ & & & . & \\ & & & & . \end{pmatrix} = \begin{pmatrix} aA & & & & \\ & bB & & & \\ & & cC & & \\ & & & . & \\ & & & & . \end{pmatrix}$$

$$UD = \begin{pmatrix} A & & & & \\ & B & & & \\ & & C & & \\ & & & . & \\ & & & & . \end{pmatrix} \begin{pmatrix} a & & & & \\ & b & & & \\ & & c & & \\ & & & . & \\ & & & & . \end{pmatrix} = \begin{pmatrix} aA & & & & \\ & bB & & & \\ & & . & & \\ & & & . & \\ & & & & . \end{pmatrix}$$

giving $aA = 1$, $bB = aA + 1 = 2$, $cC = 3$, ...

So if we try $a = 1 = A$, $b = \sqrt{2} = B$, $c = \sqrt{3} = C$, etc., then we have

$$X = U + D = \begin{pmatrix} & 1 & & & \\ 1 & & \sqrt{2} & & \\ & \sqrt{2} & & \sqrt{3} & \\ & & \sqrt{3} & & \\ & & & . & \\ & & & & . \end{pmatrix}$$

$$P = i(U - D) = \begin{pmatrix} & -i & & & \\ i & & -i\sqrt{2} & & \\ & i\sqrt{2} & & -i\sqrt{3} & \\ & & i\sqrt{3} & & \\ & & & . & \\ & & & & . \end{pmatrix}$$

and this is a solution to the commutator problem

$$[X, P]_- = 2i$$

(or $i\hbar$ if we unswap, $2 \to \hbar$).

The matrices are infinite, but they have the right property.

Let's check out the Heisenberg equations of motion

$$i\partial_t X = [X, H]_-$$

What is the Hamiltonian, H? That depends on the physical system, so let's take something familiar, the harmonic oscillator. In Note 39 of Book 8c (that's in Part IV) we found the Hamiltonian

$$H_{\rm HO} = \frac{kx^2}{2} + \frac{p^2}{2m}$$

and we worked in energy units $\hbar\omega = \hbar\sqrt{k/m}$. Let's ignore the physics of k, m, ω and $\hbar$ and focus on the essential math. So we'll pretend

$$H_{\rm HO} = \frac{1}{2}(x^2 + p^2)$$

Back to D and U

$$\begin{aligned} UD &= \frac{1}{4}(X - iP)(X + iP) \\ &= \frac{1}{4}(X^2 + P^2 + i[X, P]_-) \\ &= \frac{1}{4}(X^2 + P^2 + 2) \\ &= \frac{1}{2}(\frac{1}{2}(X^2 + P^2) + 1) \\ &= \frac{1}{2}(H_{\rm HO} + 1) \end{aligned}$$

The time rate of change of X (with $\hbar$ replaced by 2 as before, to turn the physics into math) is

$$\begin{aligned} \partial_t X &= \frac{i}{2}[H_{\rm HO}, X]_- \\ &= i[UD - \frac{1}{2}, X]_- \\ &= i[UD, U + D]_- \\ &= i(UDU - UUD + UDD - DUD) \\ &= i(U[D, U]_- + [U, D]_- D) \\ &= i(U - D) \\ &= P \end{aligned}$$

And $\partial_t P = -X$.

So $\partial_t^2 X = -X$ which is what we expect for an oscillator. The Heisenberg matrix equations of motion work with no need to take slopes.

Note, though, that this matrix approach is more abstract than the slope-equation approach: we are not drawing pictures of wavefunctions as in Notes 11, 12 and 14 above, or, for the QHO, as in

Note 39 of Book 8c.

So UD is important and worth looking at in matrix form.

$$UD = \begin{pmatrix} 0 & & & & & \\ & 1 & & & & \\ & & 2 & & & \\ & & & 3 & & \\ & & & & . & \\ & & & & & . \end{pmatrix}$$

When multiplied by energy units $\hbar\omega$ this gives the energy levels at various states of excitement of the oscillator. For instance, here are some state vectors and the corresponding energy levels: check these by multiplying by UD.

$$\begin{matrix} 0 \\ \begin{pmatrix} 1 \\ 0 \\ 0 \\ : \end{pmatrix} \end{matrix} \quad \begin{matrix} 1 \\ \begin{pmatrix} 0 \\ 1 \\ 0 \\ : \end{pmatrix} \end{matrix} \quad \begin{matrix} 2 \\ \begin{pmatrix} 0 \\ 0 \\ 1 \\ : \end{pmatrix} \end{matrix} \quad ..$$

Generally speaking the Hamiltonian will contain UD.

Note that although the energy (Hamiltonian) has discrete values 0, 1, 2, .., P and X can have continuous values. Trying to solve $Xu = \lambda u$,

$$\begin{pmatrix} & 1 & & & & \\ 1 & & \sqrt{2} & & & \\ & \sqrt{2} & & \sqrt{3} & & \\ & & \sqrt{3} & & & \\ & & & & . & \\ & & & & & . \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ : \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \\ c \\ d \\ : \end{pmatrix}$$

gives

$$\begin{aligned} b &= \lambda a \\ \sqrt{2}c &= (\lambda - 1)a \\ \sqrt{3}d &= \lambda c - \sqrt{2}b \\ &: \end{aligned}$$

which can be solved for any λ

It is not a coincidence that I've named the new operators U and D. These are *ladder* operators, going up and down, respectively, in the following sense.

$$U \mid j >= \sqrt{j+1} \mid j+1 > \quad \text{e.g.,} \quad \begin{pmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ & \sqrt{2} & 0 & & & \\ & & \sqrt{3} & 0 & & \\ & & & & . & \\ & & & & & . \end{pmatrix} \begin{matrix} j=2 \\ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ : \end{pmatrix} \end{matrix} = \sqrt{3} \begin{matrix} j+1 \\ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ : \end{pmatrix} \end{matrix}$$

$$D \mid j >= \sqrt{j} \mid j-1 > \quad \text{e.g.,} \quad \begin{pmatrix} 0 & 1 & & & & \\ & 0 & \sqrt{2} & & & \\ & & 0 & \sqrt{3} & & \\ & & & & . & \\ & & & & & . \end{pmatrix} \begin{matrix} j=2 \\ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ : \end{pmatrix} \end{matrix} = \sqrt{2} \begin{matrix} j-1 \\ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ : \end{pmatrix} \end{matrix}$$

The "ket" notation, $\mid j >$, is a convenient way of writing the jth state vector.

From this notation follows

$$UD \mid j >= \sqrt{j}U \mid j-1 >= j \mid j >$$

so UD, DU and $[D,U] = I$ are diagonal as we have seen.

Note in particular that UD is now the *number operator* N: $UD \mid j >= j \mid j >= N \mid j >$ counts the number of excitations present in state $\mid j >$.

Matrices such as U and D are going to play the respective roles of Uppear and Disappear, or create and annihilate, when we come to particles.

They will be interpreted differently, though. A particle must be in some state of energentum, which we can label, say, k. So to put a particle in this state we use the creation operator U_k. To remove it from that state we use the annihilation operator, D_k.

So to accelerate a particle from k to k' we would use

$$U_{k'} D_k$$

To annihilate a photon γ altogether and create an electron e and a positron p we might write

$$U^p_{k"} U^e_{k'} D^\gamma_k$$

with suitable constraints relating k, k' and k"—which will turn out to be impossible to satisfy in this simplistic attempt, but you get the idea. Furthermore, if these combinations of operators just give the counts of particles as the output of the math, we can multiply these numbers by energy units to get the energy and hence the Hamiltonian.

So each energentum state will have its own U and D operators. This is much more elaborate than the single U and single D for a single quantum harmonic oscillator. Each energentum state is modelled as a separate QHO—at least for bosons. The commutator relationships are

$$\begin{aligned} [D_{(-)k}, D_{(-)k'}]_- &= 0 \\ [U_{(-)k}, U_{(-)k'}]_- &= 0 \\ [D_{(-)k}, U_{(-)k'}]_- &= \delta_{kk'} \end{aligned}$$

because different states don't influence each other (the linearity of quantum mechanics), but when $k = k'$ the $D_{(-)}$ and $U_{(-)}$ don't commute but $[D_{(-)}, U_{(-)}]_- = I$.

That's for bosons, and we've explored the QHO basis already. For fermions we had anticommutators at the beginning of this Note, and it is plausible that we'll find

$$\begin{aligned} [D_{(+)k}, D_{(+)k'}]_+ &= 0 \\ [U_{(+)k}, U_{(+)k'}]_+ &= 0 \\ [D_{(+)k}, U_{(+)k'}]_+ &= \delta_{kk'} \end{aligned}$$

We're going to have to explore the matrices that have these properties. That's in the next Note.

How does all this relate to "fields"? So far we've discussed only particles. Well, how can we imagine what's going on when a "particle" is created or annihilated? It may be useful to think of it as an excitation of some sort of field.

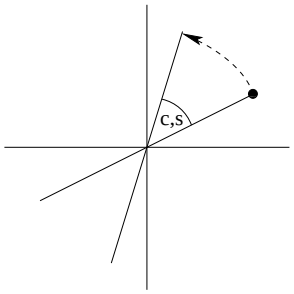
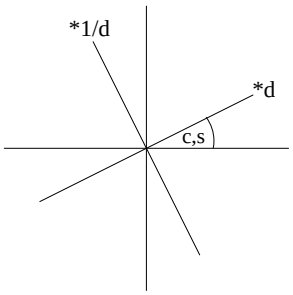
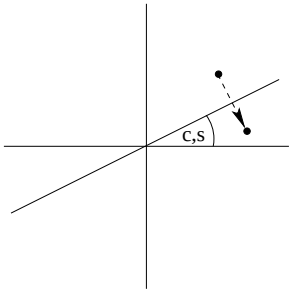
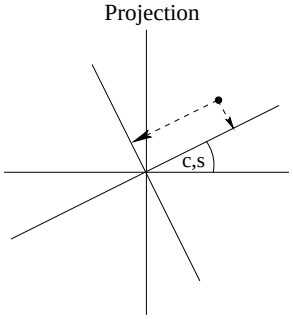
There is a precedent in crystal lattices. When they vibrate, quantum mechanics limits the levels of energy each excitation can have—they are exactly analogous to the energentum states we've just been discussing. Crystals are thought of as regularly spaced atoms connected by springs, so this is just a set of harmonic (and maybe some anharmonic) oscillators. The excitations are even

named—phonons—as if they were particles. Or, you can think of "mattress dynamics"—a regular lattice of springs—only in three dimensions.

There is a "vacuum" or null state, $| 0 >$, in which the mattress is quiescent and has no excitations. An excitation can be created by putting enough energy into the mattress to make it vibrate. The vibration can travel across the mattress, and has a certain energy (frequency) and momentum (wave number).

19. Small matrices. To explore fermion operators we don't need infinitely large matrices. Two-by-two matrices, and some logical extensions of them, will do. But we can do a lot with 2-by-2

matrices and will meet much new math. Here's a synopsis.

Category	General Form	Commutation, etc.	Squares	Cross-refs.
Rotation 	$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} =$ $cI + s\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$ $\to c + is = e^{i\theta}$ $\begin{pmatrix} c & -s \\ s & c \end{pmatrix} =$ $\begin{pmatrix} c & s \\ s & -c \end{pmatrix}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ $c^2 + s^2 = 1$	$[R, R']_{(-)} = 0$ $\det R = 1$	$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}^2 = \begin{pmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{pmatrix}$ $R_{90}^2 = -1$	Week iv Note 8 Week 4 Note 5 Week 7c Notes 7,8
Shear 	$\begin{pmatrix} c^2d + s^2/d & cs(d - 1/d) \\ cs(d - 1/d) & c^2/d + s^2d \end{pmatrix}$ $= \frac{1}{2}((d + 1/d)I +$ $(d - 1/d)\begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix}$	$[S_{c,s}, S'_{c,s}]_{(-)} = 0$ $\det S = 1$		Week iv Notes 9,10 Week 3 Note 7
Reflection 	$\begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix} =$ $c_2\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} + s_2\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ $= c_2 f_1 + s_2 f_2$	$[f_1, f_2]_{(+)} = 0$ $\det F = -1$ $\mathrm{tr} F = 0$	$F^2 = I$ $f_2 f_1 = R_{90}$	Week 7c Notes 6–11
Projection 	$P_\pm = \frac{1}{2}(I \pm F)$ e.g., $P_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $P_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	$[P_+, P_-]_{(-)} = 0$ $\det P = 0$ $\mathrm{tr} P = 1$	$P^2 = P$ $P_+ P_- = 0$	Week 7c Note 8
Ladder	$D_{(+)} = \frac{1}{2}(f_j + if_k)$ $U_{(+)} = \frac{1}{2}(f_j - if_k)$	$[D_{(+)}, U_{(+)}]_{(-)} = I$ $\det D_{(+)} = 0$ $\det U_{(+)} = 0$ $\mathrm{tr} D_{(+)} = 0$ $\mathrm{tr} U_{(+)} = 0$	$D^{+2} = 0 = U^{+2}$ e.g., $D_{(+)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ $U_{(+)} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ $D_{(+)} U_{(+)} = P_x$ $U_{(+)} D_{(+)} = P_y$	Book 8c Notes 29, 30, 39

There is a lot of information here (and some of the cross-references to these Excursions in Computing

Science are a little sketchy) so let me draw out one highlight.

Reflections are fundamental. Everything is either based on reflections or can be related to reflections. And reflections anticommute, so we should look to them, and to ladders which are based on reflections, en route to those anticommuting fermion operators.

Let's reinforce the importance of these 2-by-2 matrices by outlining the math and physics each category leads to.

Rotations lead to 2-dimensional ("complex") numbers, as indicated. They also lead to all of trigonometry: c is cosine and s is sine. Finally, I've used rotation to calculate the general forms of shear and reflection from special cases.

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} d & \\ & 1/d \end{pmatrix}\begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} c^2d+s^2/d & cs(d-1/d) \\ cs(d-1/d) & c^2/d+s^2d \end{pmatrix}$$

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}\begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix}$$

Shears exemplify symmetric matrices, that equal their transposes, and lead us to eigenvalues (e.g., d and $1/d$) which are ordinary, one-dimensional ("real") numbers for any symmetric matrix

$$S \to \begin{pmatrix} s_1 & \\ & s_2 \end{pmatrix} \qquad \det S = s_1 s_2 \qquad \text{tr } S = s_1 + s_2$$

We also have the two significant matrix invariants, that remain the same if the matrix is transformed by rotation or reflection: the determinant is the product of the eigenvalues, the trace is the sum of the eigenvalues—or of the diagonal of any matrix representing the operator with those eigenvalues.

The special case, a 45-degree shear ($c = s$)

$$S_{45} = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \qquad a^2 - b^2 = 1$$

leads to the Lorentz transformation of special relativity

$$\frac{1}{\sqrt{1-v^2}}\begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix}$$

Reflections are another example of symmetric matrices. They lead to a generalization when the matrix elements can be 2D numbers: *Hermitian matrices*. Here is a third 2-by-2 reflection

$$f_3 = \begin{pmatrix} & -i \\ i & \end{pmatrix}$$

Its square is I, it anticommutes with the other two basic reflections f_1 and f_2, its determinant is -1 and its trace is 0. It can be combined with f_1, f_2 and I, using four parameters, to give the most general Hermitian 2-by-2 matrix

$$wI + xf_2 + yf_3 + zf_1 = \begin{pmatrix} w+z & x-iy \\ x+iy & w-z \end{pmatrix}$$

The eigenvalues of Hermitian matrics are also 1D ("real") numbers, and so physically measurable. For this reason, Hermitian matrices are essential to quantum mechanics: they represent measurable quantities and contain as their eigenvalues all the possible outcomes of any measurement.

We saw an example of this in the previous Note: the quantum harmonic oscillator, QHO, has an infinity of energy levels, 0,1,2,.., all contained in the Hamiltonian matrix—which is already diagonal and so Hermitian.

The three basic reflections (or two of them) lead to the *reflection* ("Clifford") *algebra*. We can abbreviate $f_1 f_2 = f_{12}$, etc. because the subscripts and their order tell us everything we need to know.

$\times$	1	f_1	f_2	f_{12}
1	1	f_1	f_2	f_{12}
f_1	f_1	1	f_{12}	f_2
f_2	f_2	$-f_{12}$	1	$-f_1$
f_{12}	f_{12}	$-f_2$	f_1	-1

$\Rightarrow$

$[]_\pm$	1	f_1	f_2	f_{12}
1		$-$	$-$	$-$
f_1			$+$	$+$
f_2				$+$
f_{12}				

(The second table is another way of looking at the products of different elements.)

For example, $f_1 f_{12} = f_1 f_1 f_2 = f_2$ can be seen as f_{112} with the rule that adjacent same indices eliminate each other.

Or $f_2 f_{12} = f_2 f_1 f_2 = -f_1 f_2 f_2 = -f_1$ can be seen as f_{212} with the rule that swapping two not-same indices changes the sign.

Note that the double reflection f_{12} is just a 90-degree rotation, whose square is -1. Note also that f_1 is the reflection in the horizontal axis while f_2 is the reflection in the 45-degree line. Their combination

$$f_{21} = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$$

is the 90-degree rotation. This half-angle (90-degree rotation from 45-degree reflection), indicated also in the table by the difference between

$$\text{rotation} \qquad \begin{pmatrix} c & s \\ s & -c \end{pmatrix} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

and

$$\text{reflection} \qquad \begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix}$$

is going to be significant for quantum mechanics.

Ladders lead to the *anticommutator* (Grassmann) *algebra* with

$$\begin{aligned} [D_j, U_k]_+ &= \delta_{jk} \\ [D_j, D_k]_+ &= 0 \quad \text{implying } D_j^2 = 0 \\ [U_j, U_k]_+ &= 0 \quad \text{implying } U_j^2 = 0 \end{aligned}$$

which is just what we want. Unfortunately we have only one D and one U so far, so we don't have an example of the full algebra.

However, f_3 gives the nicest ladder operators and we'll use them.

$$\begin{aligned} D &= \frac{1}{2}(f_2 + i f_3) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ U &= \frac{1}{2}(f_2 - i f_3) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

20. Tensor products. To find multiple Ds and Us we'll extend the set of basic reflections. It is clear that the reflection (Clifford) algebra based on n reflections has 2^n elements. We so far have a 2-by-2 matrix representation for the 8 elements of an algebra with $n = 3$. This is as far as we can go with 2-by-2 matrices, which have $2 \times 2 \times 2 = 8$ numbers (including both parts of 2D numbers) as coefficients. For instance in our algebra based on f_1, f_2 and f_3, we had Hermitian matrices

$$wI + x f_2 + y f_3 + z f_1$$

and we can add to that the anti-Hermitian matrices ($A^\dagger = iA$)

$$w' f_{123} + x' f_{31} + y' f_{12} + z' f_{23}$$

To get bigger matrices we introduce the *tensor product*

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \overset{\leftarrow}{\times} \begin{pmatrix} e & g \\ f & h \end{pmatrix} = \begin{pmatrix} ae & ag & ce & cg \\ af & ah & cf & ch \\ be & bg & de & dg \\ bf & bh & df & dh \end{pmatrix}$$

The notation reflects the asymmetry of the operation: the second matrix is transferred four times into the first and multiplied each time by the corresponding element of the first.

Note that a complementary notation is not needed:

$$A \overset{\rightarrow}{\times} B = B \overset{\leftarrow}{\times} A$$

An important property is

$$(A \overset{\leftarrow}{\times} B)(C \overset{\leftarrow}{\times} D) = (AB) \overset{\leftarrow}{\times} (CD)$$

where adjacency means the usual matrix multiplication.

This means that we can combine 2-by-2 base reflections into a larger set of 4-by-4 reflections, and also hope to get a bigger set of ladder operations.

Let's try the following combination of f_1, f_2 and f_3. Only I'm going to rename them to be more conventional, to f_z, f_x and f_y respectively. Then, for example

$$wI + xf_x + yf_y + zf_z = \begin{pmatrix} w+z & x-iy \\ x+iy & w-z \end{pmatrix}$$

I'll also reuse the f_js as the basis "reflections" in an $n = 4$ algebra with 16 elements, starting at 0 to match the usual convention in relativity.

$$\begin{array}{cccc} f_0 & f_1 & f_2 & f_3 \\ f_z \overset{\leftarrow}{\times} I & f_x \overset{\leftarrow}{\times} f_x & f_x \overset{\leftarrow}{\times} f_y & f_x \overset{\leftarrow}{\times} f_z \end{array}$$

Clearly the last three anticommute among themselves: the products are $I \overset{\leftarrow}{\times} f_{xy}$, etc., and their commuted forms are $I \overset{\leftarrow}{\times} f_{yx} = -I \overset{\leftarrow}{\times} f_{xy}$. The first anticommutes with the last three, having the form $f_{zx} \overset{\leftarrow}{\times} f_x$, etc., added to $f_{xz} \overset{\leftarrow}{\times} f_x$, etc..

This solution is far from unique, but there are combinations that do not anticommute: you should experiment with a few yourself.

The squares are all I (which in this context stands for the 4-by-4 identity matrix).

We can now get two down and two up ladders, e.g.,

$$\begin{aligned} D_1 &= \frac{1}{2}(f_0 + if_1) & \qquad D_2 &= \frac{1}{2}(f_2 + if_3) \\ U_1 &= \frac{1}{2}(f_0 - if_1) & \qquad U_2 &= \frac{1}{2}(f_2 - if_3) \end{aligned}$$

So tensor products of reflections give us multiple down and up ladder operators which we can use to annihilate or create fermions. We won't need to work with the resulting matrices—the properties of the operators suffice—but it is nice to know that we can write matrices for them if we want.

Tensor products of matrices with themselves give results which simplify in useful ways.

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \overset{\leftarrow}{\times} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} aa & ac & ca & cc \\ ab & ad & cb & cd \\ ba & bc & da & dc \\ bb & bd & db & dd \end{pmatrix}$$

There are identical elements in the second and third rows and again in the second and third columns. If they were subtracted from each other we could get some zeros—maybe enough to block-diagonalize the 4-by-4 matrix.

Let's write a transformation which adds and subtracts these elements. Here's one which is its own inverse (it is based on reflections).

$$\begin{pmatrix} 1 & & & \\ & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \\ & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \\ & & & 1 \end{pmatrix}$$

Pre- and post-multiplying, then swapping rows 3 and 4, and columns 3 and 4, gives

$$\begin{pmatrix} a^2 & ac\sqrt{2} & c^2 & \\ ab\sqrt{2} & ad+bc & cd\sqrt{2} & \\ b^2 & bd\sqrt{2} & d^2 & \\ & & & ad-bc \end{pmatrix}$$

Note that the singleton entry, $ad - bc$ is the determinant of the original 2-by-2, and could be something simple such as 1, -1 or 0.

We can take the 3-by-3 block and combine it with the original matrix again to get a 6-by-6. The top row is

$$a^3 \quad a^2c \quad a^2c\sqrt{2} \quad ac^2\sqrt{2} \quad ac^2 \quad c^3$$

and similar relationships in the last row and in the first and last columns.

This suggests a transformation containing blocks

$$\begin{matrix} \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} \\ \frac{\sqrt{2}}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{matrix} \quad \text{and} \quad \begin{matrix} \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \end{matrix}$$

which, when applied and the result rearranged, gives a 4-by-4 block

$$\begin{matrix} a^3 & a^2c\sqrt{3} & ac^2\sqrt{3} & c^3 \\ a^2b\sqrt{3} & a^2d+2abc & bc^2+2acd & c^2d\sqrt{3} \\ ab^2\sqrt{3} & b^2c+2abd & ad^2+2bcd & cd^2\sqrt{3} \\ b^3 & b^2d\sqrt{3} & bd^2\sqrt{3} & d^3 \end{matrix}$$

and a 2-by-2 block which is $ad - bc$ (that determinant again) times the original 2-by-2 matrix.

There are two ways we can take the next step to a 5-by-5: either take the tensor product of the 4-by-4 above with the original 2-by-2 and block diagonalize

$$4 \otimes 2 = 5 \oplus 3$$

or use the tensor product of the earlier 3-by-3 with itself

$$3 \otimes 3 = 5 \oplus 3 \oplus 1$$

In each case, the "5"s are the same, the "3"s are $ad - bc$ times the "3" we first generated, and the "1" is $(ad - bc)^2$.

The notation I've used, with the circles and the $\times$ and $+$, just keeps track of the dimensions of the spaces. Thus when we tensor-product two 3-D spaces the result decomposes into blocks of 5-D, 3-D and 1-D spaces.

21. Spin. The shear and reflection categories in the big table at the beginning of Note 19 both have angle (c_2, s_2)—indeed the role this angle plays in shear is exactly the reflection matrix.

The significance is that if we rotate a two-sided mirror (one which reflects both ways, which is what the matrix describes) through 180^o we've returned the system to its starting point. This is also true of any shear.[1]

A system which is symmetrical in this way under half a full rotation can be described as having spin 2.

By contrast, a system with no particular symmetry must be rotated all the way around, and would have spin 1.

A system with *two* mirrors, though, generates a rotation. The one we showed had a mirror $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ along the horizontal axis, and a mirror $\begin{pmatrix} c & s \\ s & -c \end{pmatrix}$ at *half* the angle (c, s) from the horizontal one: note that the reflection from a mirror at angle (c, s) is described by c_2 and s_2.

So in principle to return the 2-mirror system to its unchanged starting point would require a 720-degree rotation of the second mirror. This is a little hard to visualize for mirrors but the math is clear.

Such a system, which needs a double rotation to restore it, has spin 1/2.

This situation is beyond our everyday experience. But the math says reflections are fundamental and, in particular, a rotation by (c_2, s_2) is two reflections separated by half the angle, (c, s). One might believe that if the math says something is fundamental then we will see it in nature. And we do: spin 1/2 is essential for fermions, An Excursion in Week 7a lets you work through Pauli's reasoning that a relativistic observer cannot distinguish swapping two fermions from rotating one fermion through 360 degrees. In both case the sign of the amplitude changes. This sign only comes back to $+1$ if you do 360^o *twice*.

Half-angles will be sufficiently important that they need notation. Since capital "J" looks sort of like an upside-down "2", I'll take that as the subscript indicating the half angle—c_J and s_J—just as c_2 and s_2 indicates the double angle.

$$\begin{pmatrix} c_J & -s_J \\ s_J & c_J \end{pmatrix}^2 = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \qquad \begin{pmatrix} c & -s \\ s & c \end{pmatrix}^2 = \begin{pmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{pmatrix}$$

Spin 1/2 is probably goiing to be fundamental. Let's approach it via spin 1 and spin 2 in three dimensions.

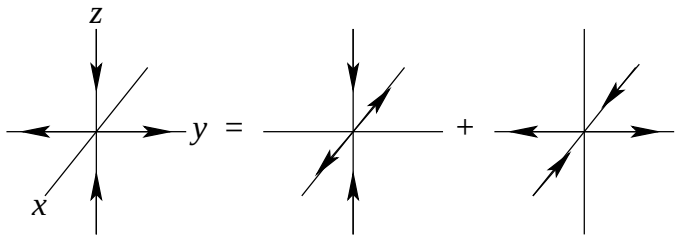
Spin 1 is just ordinary rotation and in 3D there are three independent directions we can rotate in. So spin 1 will be represented by a 3-by-3 matrix.

For spin 2 in 3D we must think about 2-sided mirrors or about shears. Indeed, gravity waves, predicted by general relativity and detected directly in 2015 a century after Einstein published his

[1] Note that the two independent reflections are 45^o apart, not 90^o.

first paper on GR, are shear waves and have spin 2.

Here are three shears along the three pairs of axes possible in 3D, matrix representations of them, and their relationships.



$$\begin{pmatrix} 0 & & \\ & \epsilon & \\ & & -\epsilon \end{pmatrix} = \begin{pmatrix} \epsilon & & \\ & 0 & \\ & & -\epsilon \end{pmatrix} + \begin{pmatrix} -\epsilon & & \\ & \epsilon & \\ & & 0 \end{pmatrix}$$

We've had to use shear *generators* rather than matrices representing the shears themselves. If we let

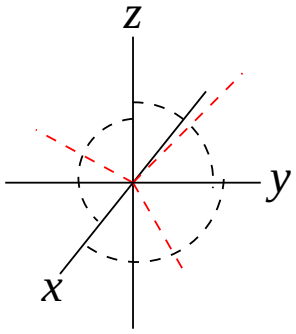
$$d \approx 1 + \epsilon \quad \text{and} \quad \frac{1}{d} \approx 1 - \epsilon$$

for very small displacements ϵ, then, say

$$\begin{pmatrix} 1 & & \\ & d & \\ & & 1/d \end{pmatrix} = I + \begin{pmatrix} 0 & & \\ & \epsilon & \\ & & -\epsilon \end{pmatrix} = I + \epsilon \begin{pmatrix} 0 & & \\ & 1 & \\ & & -1 \end{pmatrix}$$

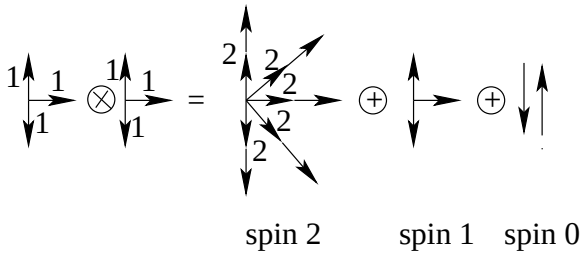
The upshot of the above combination of the three shears is that only two of them are independent.

In addition, though, there are three shears at 45^o between each pair of axes and they re all independent of each other and of the previous two.

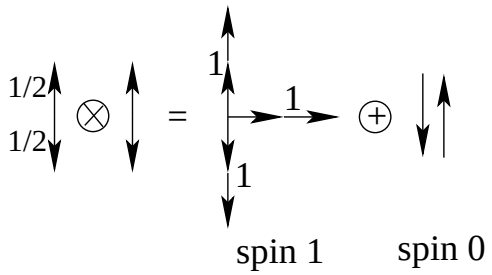


So spin 1 has three independent directions in 3D and spin 2 has five independent directions.

This might remind us of how we combined two 3-by-3 matrices in Note 20 to get a 5-by-5 plus change: $3 \otimes 3 = 5 \oplus 3 \oplus 1$. We might draw these combinations schematically and that would give us a way of visualizing the 5 dimensions of spin 2, the 3 dimensions of spin 1 and the 1 dimension of spin 0..



It also suggests 2 dimensions for spin 1/2, showing the first combination we did in Note 19 as



So we must look for a 2-dimensional representation of 3D rotations. (And a 3-dimensional representation, which will be trivial, .., and a 5-dimensional representation, etc.).

To do this we'll consider *rotation generators* and find some interesting math. First 2D rotations.

A very small rotation is

$$R_\theta = \begin{pmatrix} c_\theta & -s_\theta \\ s_\theta & c_\theta \end{pmatrix} \theta \overrightarrow{\to} 0 \begin{pmatrix} 1 & -\theta \\ \theta & 1 \end{pmatrix} = I - i\theta J$$

with

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

(We don't need the two is, which oppose each other after all, but it is conventional to have them.) Now consider, since $J^2 = I$,

$$\begin{aligned} e^{-i\theta J} &= I - i\theta J + \frac{1}{2}(-i\theta J)^2 - \frac{1}{3!}(-i\theta J)^3 + .. \\ &= I(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - ..) \\ &\quad -iJ(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - ..) \\ &= I\cos\theta - iJ\sin\theta \\ &= \begin{pmatrix} c_\theta & -s_\theta \\ s_\theta & c_\theta \end{pmatrix} \end{aligned}$$

So now for any angle θ, given the generator J, we can get back the rotation

$$R_\theta = e^{-i\theta J}$$

This works for 3D rotations which, together with their corresponding generators, are

$$\begin{array}{ccc} R_x & R_y & R_z \\ \begin{pmatrix} 1 & & \\ & c & -s \\ & s & c \end{pmatrix} & \begin{pmatrix} c & & s \\ & 1 & \\ -s & & c \end{pmatrix} & \begin{pmatrix} c & -s & \\ s & c & \\ & & 1 \end{pmatrix} \\ J'_x & J'_y & J'_z \\ \begin{pmatrix} 0 & & \\ & 0 & -i \\ & i & 0 \end{pmatrix} & \begin{pmatrix} 0 & & i \\ & 0 & \\ -i & & 0 \end{pmatrix} & \begin{pmatrix} 0 & -i & \\ i & 0 & \\ & & 0 \end{pmatrix} \end{array}$$

It will be instructive to transform all of these matrices so that one of them—say J_z—is diagonal. Say $J_z = XJ'_zX^{-1}$. Try

$$X = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i & \\ & & -\sqrt{2} \\ -1 & -i & \end{pmatrix}$$

and confirm that X^{-1} is the Hermitian conjugate, $X^\dagger = X^{-1}$.

The results, $J_x = XJ'_xX^{-1}$, $J_y = XJ'_yX^{-1}$ and $J_z = XJ'_zX^{-1}$, are

$$\begin{matrix} J_x & & J_y & & J_z \\ \frac{1}{\sqrt{2}}\begin{pmatrix} & 1 & \\ 1 & & 1 \\ & 1 & \end{pmatrix} & \frac{1}{\sqrt{2}}\begin{pmatrix} & -i & \\ i & & -i \\ & i & \end{pmatrix} & \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix} \end{matrix}$$

These eigenvalues, 1, 0 and -1 correspond nicely to the symbolic pictures we drew for spin 1 above in this Note. They are also the eigenvalues for J_x and J_y and we could have diagonalized either of them instead of J_z.

Note that no two of the original J'_x, J'_y and J'_z can be diagonalized simultaneously. This means that no two of them commute. Indeed, the commutators of these Js are very significant.

$$\begin{aligned} [J_x, J_y]_- &= iJ_z \\ [J_y, J_z]_- &= iJ_x \\ [J_z, J_x]_- &= iJ_y \end{aligned}$$

$[]_-$	J_x	J_y	J_z
J_x		iJ_z	$-iJ_y$
J_y			iJ_x
J_z			

But we decided that spin 1/2 would need 2-by-2 matrices. Can we find 2-by-2 matrices with the same commutators?

Analogy helps. How about

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

This is great. We've seen these matrices before. They are 1/2 of the basic 2-by-2 reflections. They satisfy an algebra, the *commutator* (Lie) algebra

$$[J_j, J_k]_- = \sum_\ell \ell_{jk\ell} J_\ell$$

where in this case the coefficient $\ell_{jk\ell} = i\epsilon_{jk\ell}$, being i if j, k, ℓ are an even permutation of 1,2,3, $-i$ if the permutation is odd, and 0 otherwise.

The eigenvalues of each of these matrices (and of any combination $pJ_x + qJ_y + rJ_z$ with $p^2+q^2+r^2 = 1$) are 1/2 and $-1/2$. These also match the symbolic pictures we drew earlier in this Note for spin 1/2.

Because $J_k = f_k/2$ and $f_k^2 = I$ for the 2-by-2 matrices, $k = x, y, z$, we can expand

$$e^{-i\theta_k J_k} = e^{-i(\theta_k/2)f_k} = I\cos(\theta_k/2) - if_k \sin(\theta_k/2)$$

Note the half angles!

So

$$\begin{aligned} e^{-i\theta_x J_x} &= \begin{pmatrix} c_x & -is_x \\ -is_x & c_x \end{pmatrix} \\ e^{-i\theta_y J_y} &= \begin{pmatrix} c_y & -s_y \\ s_y & c_y \end{pmatrix} \\ e^{-i\theta_z J_z} &= \begin{pmatrix} c_z - is_z & \\ & c_z + is_z \end{pmatrix} = \begin{pmatrix} e^{-i\theta_z/2} & \\ & e^{i\theta_z/2} \end{pmatrix} \end{aligned}$$

where, here, $c_x = \cos(\theta_x/2)$, $s_x = \sin(\theta_x/2)$, and similarly for y and z.

With these we can find the 2-by-2 matrix for any 3D rotation. Often, to go from one axis system, x, y, z, to another, x', y', z', the *Euler angles* are used: rotate by angle β about the z-axis, then by angle α about the x-axis, then finally by angle γ about the z-axis (which is of course no longer the original z-axis).

$$\begin{pmatrix} e^{-i\gamma/2} & \\ & e^{i\gamma/2} \end{pmatrix} \begin{pmatrix} c_{\alpha/2} & -is_{\alpha/2} \\ -is_{\alpha/2} & c_{\alpha/2} \end{pmatrix} \begin{pmatrix} e^{-i\beta/2} & \\ & e^{i\beta/2} \end{pmatrix}$$

$$= \begin{pmatrix} c_{\alpha/2}e^{-i(\beta+\gamma)/2} & -is_{\alpha/2}e^{i(\beta-\gamma)/2} \\ -is_{\alpha/2}e^{-i(\beta-\gamma)/2} & c_{\alpha/2}e^{i(\beta+\gamma)/2} \end{pmatrix}$$

We can now use the tensor product

$$2 \otimes 2 = 3 \oplus 1$$

to go from the spin-1/2 rotation above to the spin-1 rotation, but I did that and discussed it in Note 10 of Week 6.

We can also use ladder operators

$$\begin{aligned} J_+ &= \frac{1}{\sqrt{2}}(J_x + iJ_y) \\ J_- &= \frac{1}{\sqrt{2}}(J_x - iJ_y) \end{aligned}$$

to construct the J matrices in higher dimensions, supposing that J_z follows the patterns established above

$$J_z = \begin{pmatrix} 1/2 & \\ & -1/2 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} 3/2 & & & \\ & 1/2 & & \\ & & -1/2 & \\ & & & -3/2 \end{pmatrix}$$

But I did that in Note 29 of Book 8c (in Part III).

22. Vectors and spinors. In practice, we are interested in 2D, 3D and 4D space, so here is the 16-element full reflection algebra[2] for 4D. I've shown the squares down the diagonal and, above the diagonal, whether the products commute ($-$) or anticommute ($+$). Below the diagonal is the special case of eight elements in 3D (but without f_0). You can extract from this the four 2D elements (no f_3).

[2]The usual name honours William Kingdon Clifford, 1845–79

I	I	I	f_0	f_1	f_2	f_3	f_{01}	f_{02}	f_{03}	f_{23}	f_{31}	f_{12}	f_{123}	f_{023}	f_{013}	f_{012}	f_{0123}
		I	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$
d_0	f_0		I	$+$	$+$	$+$	$+$	$+$	$+$	$-$	$-$	$-$	$+$	$-$	$-$	$-$	$+$
d_{01}	f_1	$-$		I	$+$	$+$	$+$	$-$	$-$	$-$	$+$	$+$	$-$	$+$	$-$	$-$	$+$
d_{02}	f_2	$-$		$+$	I	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$-$	$+$	$-$	$+$
d_{03}	f_3	$-$		$+$	$+$	I	$-$	$-$	$+$	$+$	$+$	$-$	$-$	$-$	$-$	$+$	$+$
d_1	f_{01}						$-I$	$+$	$+$	$-$	$+$	$+$	$+$	$+$	$-$	$-$	$-$
d_2	f_{02}							$-I$	$+$	$+$	$-$	$+$	$+$	$-$	$+$	$-$	$-$
d_3	f_{03}								$-I$	$+$	$+$	$-$	$+$	$-$	$-$	$+$	$-$
$-d_{23}$	f_{23}	$-$		$-$	$+$	$+$				$-I$	$+$	$+$	$-$	$-$	$+$	$+$	$-$
$-d_{31}$	f_{31}	$-$		$+$	$-$	$+$				$+$	$-I$	$+$	$-$	$+$	$-$	$+$	$-$
$-d_{12}$	f_{12}	$-$		$+$	$+$	$-$				$+$	$+$	$-I$	$-$	$+$	$+$	$-$	$-$
$-d_{0123}$	f_{123}	$-$		$-$	$-$	$-$				$-$	$-$	$-$	$-I$	$+$	$+$	$+$	$+$
$-d_{023}$	f_{023}													$-I$	$+$	$+$	$+$
$-d_{013}$	f_{013}														$-I$	$+$	$+$
$-d_{012}$	f_{012}															$-I$	$+$
$-d_{123}$	f_{0123}																I

I've also written a column of new names and called them $d_0, ...,$ for Dirac. Here's what we can do with the fs in 2D and 3D Euclidean space and with the ds in 4D Minkowski space.

In 3D, given a vector $\vec{a} = (a_1, a_2, a_3)$, we can write $\not{a} = a_1 f_1 + a_2 f_2 + a_3 f_3$. Then the Gibbsian dot product between $\vec{a}$ and $\vec{b}$ is

$$\vec{a} \cdot \vec{b} = \frac{1}{2}[\not{a}\ ,\not{b}\ ]_+ = a_1 b_1 + a_2 b_2 + a_3 b_3$$

while the cross product is, effectively

$$\vec{a} \times \vec{b} = \frac{1}{2}[\not{a}\ ,\not{b}\ ]_-$$

In particular,

$$\begin{aligned} \not{a}\,\not{a} &= (a_1 f_1 + a_2 f_2 + a_3 f_3)(a_1 f_1 + a_2 f_2 + a_3 f_3) \\ &= a_1^2 + a_2^2 + a_3^2 \\ &= \vec{a}^2 \end{aligned}$$

This works in 2D too. Let's look at

$$\frac{1}{2}[\not{a}\ ,\not{b}\ ]_- = (a_1 b_2 - a_2 b_1) f_{12}$$

We interpret f_{12} as the *xy plane*, and we could think of it as a vector perpendicular to that plane if we want. That's the Gibbs interpretation in 3D.

We could extend this notation to 4D, but for relativity and 4D Minkowski space we find we should use the "Dirac" symbols, d, from the above table instead of the reflections f. We write

$$\not{a} = a_0 d_0 + a_1 d_1 + a_2 d_2 + a_3 d_3$$

Again,

$$\frac{1}{2}[\not{a}, \not{b}]_+ = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3$$

and this has the right signs for the invariant scalar product of two 4-vectors. Again in particular

$$\not{a}\ \not{a} = a_0^2 - a_1^2 - a_2^2 - a_3^2 = a^2$$

for Minkowski space.

We can define the "dot product" for both Euclidean 3D and Minkowski 4D

$$a \cdot b \stackrel{\text{def}}{=} \frac{1}{2}[a, b]_+$$

The commutator generalizes the cross product to 4D but is not used

$$\begin{aligned}\frac{1}{2}[\not{a}, \not{b}]_- &= (a_0b_1 - b_0a_1)f_1 - (a_2b_3 - a_3b_2)f_{23} \\ &\quad +(a_0b_2 - b_0a_2)f_2 + (a_1b_3 - a_3b_1)f_{31} \\ &\quad +(a_0b_3 - b_0a_3)f_3 - (a_1b_2 - a_2b_1)f_{12}\end{aligned}$$

Just as the backslash notation, $\backslash\!\!\!a$, releases us from limitation to 3D (and 2D) of Gibbs vectors, so the slash notation, $\not{a}$, releases us from needing the Minkowski metric, $\texttt{diag}(1, -1, -1, -1)$, and the consequent distinction between co- and contravariant indices. We must make one distinction: although a^μ, say, represents (a_0, a_1, a_2, a_3), for *slopes* ∂^μ represents $(\partial_t, -\partial_x, -\partial_y, -\partial_z)$. This is because, in the up, down notations of contra-, co-variance, $\partial^\mu = \partial/\partial x_\mu$ and $\partial_\mu = \partial/\partial x^\mu$, so, in that notation a_μ is $(a_0, -a_1, -a_2, -a_3)$, but ∂_μ is $(\partial_t, \partial_x, \partial_y, \partial_z)$.

So slopes are opposite. The net effect is

$$\begin{aligned}\frac{1}{2}[\not{a}, \not{b}]_+ &= a_0b_0 - \vec{a} \cdot \vec{b} \\ \frac{1}{2}[\not{\partial}, \not{a}]_+ &= \partial_t a_0 + \vec{\bigtriangledown} \cdot \vec{a} \\ \not{a}\ \not{a} &= a_0^2 - \mid \vec{a} \mid^2 \\ \not{\partial}\ \not{\partial} &= \partial_t^2 - \bigtriangledown^2\end{aligned}$$

with conventional meanings, $\vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + a_3b_3$, $\mid \vec{a} \mid^2 = a_1^2 + a_2^2 + a_3^2$ and $\bigtriangledown = (\partial_x, \partial_y, \partial_z)$. (That is, $\vec{a} \cdot \vec{b} = \frac{1}{2}[\backslash\!\!\!a\ , \backslash\!\!\!b\ ]_+$, $\mid \vec{a} \mid^2 = \vec{a} \cdot \vec{a} = \backslash\!\!\!a\ \backslash\!\!\!a$ and $\vec{\bigtriangledown} \cdot \vec{a} = \frac{1}{2}[\backslash\!\!\!\partial\ , \backslash\!\!\!a\ ]_+$).

In Note 9 of Week 7c we exploited the fact that a rotation is two reflections to write a rotation in a 3D plane, P, of a vector v by and angle (c, s) as

$$(c_J - s_J P)v(c_J + s_J P)$$

For instance (the general result is in Week 7c)

$$\begin{aligned}v &= yf_2 + zf_3 \\ P &= f_{23} \\ (c_J - s_J f_{23})(yf_2 + zf_3)(c_J + s_J f_{23}) &= (f_2\ f_3)\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}\end{aligned}$$

which gives the 2D special case.

This also works in 4D and in 4D Minkowski space using the d elements, e.g.,

$$\begin{aligned}v &= yd_2 + zd_3 \\ P &= d_{23} \\ (c_J - s_J d_{23})(yd_2 + zd_3)(c_J + s_J d_{23}) &= (d_2\ d_3)\begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}\end{aligned}$$

Is the reversal of sign surprising?

The result in all cases involves only the full angle, (c, s), of rotation, even though the operation, because of reflection, is written using the half angle, (c_J, s_J).

The term, $c_J - s_J f_{23}$, is analogous to the 2D-number representation of a rotation, $c + is$, and this is also another way to represent a vector.

But because of the half angle it does not return to itself on full rotation of 360^o, but to the negative of itself. So it is called a *spinor*—the reference being of course to spin 1/2.

Note that, since f_{23}, d_{23} and any special plane P square to -1, $(c_J - s_J P)(c_J + s_J P) = c_J^2 + s_J^2 = 1$ and the two spinors are inverses of each other. The rotation of a vector v given by a spinor S is SvS^{-1}.

Since the reflection algebra has matrix representations—essentially unique in 3D but varied in 4D—vectors and spinors become matrices in this approach.

In 3D, the vector

$$\not{a} = \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix}$$

and the spinor, say $c_J + s_J f_{23}$ is

$$c_J + s_J f_{23} = \begin{pmatrix} c_J & is_J \\ is_J & c_J \end{pmatrix}$$

because

$$f_{23} = f_2 f_3 = \begin{pmatrix} & -i \\ i & \end{pmatrix} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} = \begin{pmatrix} & i \\ i & \end{pmatrix}$$

For 4D we use the Dirac representation in Note 20.

$$\begin{array}{cccc} f_0 & f_1 & f_2 & f_3 \\ f_3 \overleftarrow{\times} I & f_1 \overleftarrow{\times} f_1 & f_1 \overleftarrow{\times} f_2 & f_1 \overleftarrow{\times} f_3 \\ d_0 & d_1 & d_2 & d_3 \\ f_3 \overleftarrow{\times} I & f_{31} \overleftarrow{\times} f_1 & f_{31} \overleftarrow{\times} f_2 & f_{31} \overleftarrow{\times} f_3 \end{array}$$

Here's the 4-vector.

$$\begin{aligned} \not{a} &= a_0 d_0 + a_1 d_1 + a_2 d_2 + a_3 d_3 \\ &= \begin{pmatrix} a_0 & & a_3 & a_1 - ia_2 \\ & a_0 & a_1 + ia_2 & -a_3 \\ -a_2 & -a_1 + ia_2 & & \\ -ia_1 - ia_2 & a_3 & & \end{pmatrix} \end{aligned}$$

and here's the spinor $c_J + s_J d_{23}$ with

$$d_{23} = -f_{23} = -(f_1 \overleftarrow{\times} f_2)(f_1 \overleftarrow{\times} f_3) = -I \overleftarrow{\times} f_{23}$$

(the contexts make clear when "f_{23}" means a 4-by-4 and when it means a 2-by-2).

$$\begin{pmatrix} c_J & -is_J & & \\ -is_J & c_J & & \\ & & c_J & -is_J \\ & & -is_J & c_J \end{pmatrix}$$

23. Multiple and independent systems. In Note 18 we worked out a system of two matrices, $D_{(-)}$ and $U_{(-)}$ such that

$$[D_{(-)}, D_{(-)}]_- = I$$

We have not addressed how to move from this single system to a multiple system in which

$$\begin{aligned}\left[D_{(-)j}, D_{(-)k}\right]_- &= 0\\ \left[U_{(-)j}, U_{(-)k}\right]_- &= 0\\ \left[D_{(-)j}, U_{(-)k}\right]_- &= \delta_{jk}\end{aligned}$$

as we indicated in Note 18 that we would need. The tensor product of Note 20 is the way. Here are two sets, $j, k = 1, 2$

$$\begin{aligned}D_{(-)1} &= D_{(-)} \overleftarrow{\times} I \quad U_{(-)1} = U_{(-)} \overleftarrow{\times} I\\ D_{(-)2} &= I \overleftarrow{\times} D_{(-)} \quad U_{(-)2} = I \overleftarrow{\times} U_{(-)}\end{aligned}$$

To demonstrate the commutation relationships we need a second property of the tensor product

$$A \overleftarrow{\times} C + B \overleftarrow{\times} C = (A + B) \overleftarrow{\times} C$$

and vice versa

$$C \overleftarrow{\times} A + C \overleftarrow{\times} B = C \overleftarrow{\times} (A + B)$$

From these we have

$$[A \overleftarrow{\times} C, B \overleftarrow{\times} C]_- = [A, B]_- \overleftarrow{\times} C$$

and vice versa. (And also for anticommutators.)

So

$$\begin{aligned}[D_{(-)1}, U_{(-)1}]_- &= [D_{(-)}, U_{(-)}]_- \overleftarrow{\times} I = I \overleftarrow{\times} I\\ [D_{(-)2}, U_{(-)2}]_- &= I \overleftarrow{\times} [D_{(-)}, U_{(-)}]_- = I \overleftarrow{\times} I\\ [D_{(-)1}, U_{(-)2}]_- &= (D_{(-)} \overleftarrow{\times} I)(I \overleftarrow{\times} U_{(-)}) - (I \overleftarrow{\times} U_{(-)})(D_{(-)} \overleftarrow{\times} I)\\ &= D_{(-)} \overleftarrow{\times} U_{(-)} - D_{(-)} \overleftarrow{\times} U_{(-)}\\ &= 0\end{aligned}$$

and similarly for $[D_{(-)2}, U_{(-)1}]_-$.

Thus we can easily go from $D_{(-)}$, $U_{(-)}$ to an arbitrary number of $D_{(-)j}$, $U_{(-)j}$ for $j = 1, .., n$.

$$\begin{aligned}&\quad 1 \quad 2 \qquad\qquad j \qquad\qquad n\\ D_{(-)j} &= I \overleftarrow{\times} I \overleftarrow{\times} .. \overleftarrow{\times} I \overleftarrow{\times} D_{(-)} \overleftarrow{\times} I \overleftarrow{\times} .. \overleftarrow{\times} I\\ U_{(-)j} &= I \overleftarrow{\times} I \overleftarrow{\times} .. \overleftarrow{\times} I \overleftarrow{\times} U_{(-)} \overleftarrow{\times} I \overleftarrow{\times} .. \overleftarrow{\times} I\end{aligned}$$

The composite operators are a kind of vector with the basic operators D and U appearing in ony one place.

Fortunately we don't have to write all this out, and especially we don't have to multiply out the tensor products. This is because a combination such as (the repeated js are *not* summed)

$$\begin{aligned}A_j b_j &= (I \overleftarrow{\times} .. \overleftarrow{\times} A \overleftarrow{\times} .. \overleftarrow{\times} I)(1 \overleftarrow{\times} .. \overleftarrow{\times} b \overleftarrow{\times} .. \overleftarrow{\times} 1)\\ &= 1 \overleftarrow{\times} .. \overleftarrow{\times} Ab \overleftarrow{\times} .. \overleftarrow{\times} 1\end{aligned}$$

where, if b is a matrix, "1" is just I, but if b is a vector, "1" is the same-size vector of all 1s.

The result is that only the jth components are affected and we could write, simply,

$$A_j b_j = (Ab)_j$$

Then a mix would sort itself out automatically.

$$A_1 A_2 b_1 b_2 = (A_1 b_1)(A_2 b_2)$$

So we need only label these commuting operators to distinguish them from one another, and write them in sequence.

This is a fortunate simplification, especially since the basic $D_{(-)}$ and $U_{(-)}$ of Note 18 are themselves infinite matrices.

The original $D_{(-)}$ and $U_{(-)}$ described a quantum harmonic oscillator, QHO. We are now seeing a boson field as a collection of QHOs, one for each energentum state (*not* one for each boson, in case that may be a confusion). Instead of counting energy levels in the single QHO, the number operator $U_{(-)}D_{(-)}$ must now be interpreted as counting the number of particles in the state described by the QHO. For bosons, any number of particles can occupy the same state—the more the merrier.

While we are simplifying notation, boson states in particular need simplifying. A QHO excited to its nth energy level, or a boson state containing n bosons, would be written as an infinite vector of 0s with a single 1, in the $n+1$st position (we must permit $n = 0$). It is more compact to replace this by

$$| \, n >$$

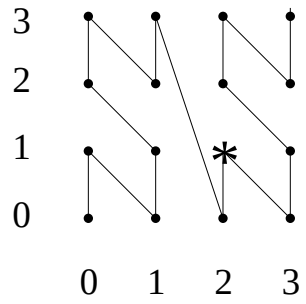
Then the number of bosons in each of many independent states, instead of being

$$| \, n_1 > \overleftarrow{\times} | \, n_2 > \overleftarrow{\times} \, .. \, \overleftarrow{\times} | \, n_k > \overleftarrow{\times} \, ..$$

would be

$$| \, n_1 n_2 .. n_k .. >$$

Both notations presuppose a way of enumerating all the independent states. With discrete energy and momentum values, all states might form a 4D lattice (I show a 2D example below). Then some variant of "Z-order" might be used. Here is a 2D example.



The position in the ordering is given by interleaving the bits of the coordinates. Thus, the position of, say, (2,1)—starred in the diagram—is

$$(2,1) = (1_2 0_1, 0_4 1_3) \rightarrow 1_2 0_4 0_1 1_3 = 9$$

(I've labelled each bit to show the interleaving. The least significant bits are on the right for each number.)

But we will work with examples too simple to need such generalization.

Note that two such "kets" are orthogonal unless identical.

$$| \, n_1 n_2 .. n_k .. > \perp | \, n'_1 n'_2 .. n'_k .. >$$

if $n_k \neq n'_k$ for any k.

The way to express this as a form of inner product is to cast one as a "bra" (thus "bra|ket")

$$< n_1 n_2 .. n_k .. \mid n'_1 n'_2 .. n'_k .. >= \prod \delta_{n_k n'_k}$$

and this also shows that the bras and the kets are normalized.

We are also going to have to mix up $D_{(-)}$ and $U_{(-)}$ for bosons with the finite $D_{(+)}$ and $U_{(+)}$ for fermions.

We showed in Note 21 (well, it is an Excursion to Note 21) that two matrices are simultaneously diagonalizable—i.e., by the same transformation—if and only if they commute.

Since Heisenberg quantum mechanics interprets (Hermitian) matrices as quantities and their eigenvalues as possible outcomes of measurements, we must suppose that *independent* quantities can be measured simultaneously and so that they commute with each other as matrices.

Hence the multiple $D_{(-)}$s and $U_{(-)}$s we've just discussed can either describe different states of a single boson (different energentum values, for instance) or states of completely different bosons, say a photon and an electroweak W boson.

When we come to fermions, the multiple $D_{(+)}$ and $U_{(+)}$ operators may apply to different states of the same fermion, in which case they anticommute.

$$\begin{aligned} \left[D_{(+)j}, D_{(+)k}\right]_+ &= 0 \\ \left[U_{(+)j}, U_{(+)k}\right]_+ &= 0 \\ \left[D_{(+)j}, U_{(+)k}\right]_+ &= \delta_{jk} \end{aligned}$$

or they may apply to different fermions, in which case everything commutes.

The former situation we addressed in Note 20, using tensor products to extend the 2^3-element reflection algebra, with one ladder each way ($D_{(+)}$ and $U_{(+)}$) to the 2^4-element reflection algebra with two ladders, and so on.

Independence of multiple sets of such ladder systems can be generated by combining them into commutative tensor products just as we did for boson ladders $D_{(-)}$ and $U_{(-)}$ above.

The abstract space in which these ladders reside and can be distinguished from each other is called a Fock space.

24. A simple field. Before Schrödinger wrote down his equation, based on

$$E = \frac{p^2}{2m} + V$$

and the wave equivalents (Note 10)

$$\begin{aligned} E = \hbar\omega &\rightarrow i\hbar\partial_t \\ p_j = \hbar k_j &\rightarrow -i\hbar\partial_{x_j} \end{aligned}$$

he considered the relativistic equation

$$E^2 - P^s c^2 = m^2 c^4$$

and the same wave equivalents.

Using notation from Note 22, and working in units with $c = 1$ and $\hbar = 1$ (see Excursion *Waves* in Week 7c) this becomes $\omega^2 - k^2 = m^2$ and (the "Klein-Gordon equation")

$$(\not{\partial} \cdot \not{\partial} + m^2)\phi = 0$$

since

$$\not{\partial}\cdot\not{\partial} = \frac{1}{2}[\not{\partial},\not{\partial}]_+ = \partial_t^2 - \partial_x^2 - \partial_y^2 - \partial_z^2$$

(and don't forget the i in the transitions from E and p to ∂.)

A solution to this is

$$\phi = \frac{1}{\sqrt{L^3}} e^{-i\not{k}\cdot\not{x}}$$

with

$$\not{k}\cdot\not{x} = \frac{1}{2}[\not{k},\not{x}]_+ = \omega_k t - k_x x - k_y y - k_z z$$

and the multiplying constant involving the volume of a cubical box with sides of length L, which we'll come to later. (I've written the k explicitly in ω_k because $\omega_k^2 = k^2 + m^2$ is a function of k.)

We could interpret this $\phi(t,x,y,z)$ as a wave function but doing so gave Schrödinger sufficient problems that he abandoned it for the wavefunction of the Schrödinger equation (Note 10).

Besides, the wavefunction interpretation does not allow creation or annihilation of states, which we found in Note 18 to be important.

Instead we will interpret $\phi(t,x,y,z)$ as a *field*, the excitations of which will be particles which we can create and annihilate. The creation and annihilation operators will work in momentum-space, or k-space, and will add and remove particles of a specified momentum. So we will be needing Fourier transforms (Week 9) to map from energentum space (ω, k_x, k_y, k_z) to timespace $(t,x,y.z)$.

As we discussed in Book 8c, Notes 36 and 37 (Part IV of Book 8c) calculations are simplified if we can capture all the physics in a simple function, rather than working with vectors, etc. (of, say, force and momentum). The simple functions we considered there started with potential energy, $V(x,y,z)$, and advanced to the Lagrangian, L, and the Hamiltonian, H.

$$\begin{pmatrix} L \\ H \end{pmatrix} = \begin{pmatrix} T-V \\ T+V \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} T \\ V \end{pmatrix} \qquad \begin{pmatrix} T \\ V \end{pmatrix} = \frac{1}{2}\begin{pmatrix} H+L \\ H-L \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} L \\ H \end{pmatrix}$$

with T being the kinetic energy.

In Notes 38 and 39 of Book 8c we found a direct connection between Lagrangian and Hamiltonian via "generalized coordinates" q (generalized position) and p (generalized momentum) and a "Legendre transformation"

$$H = p\dot{q} - L$$

with p defined

$$p = \partial_{\dot{q}} L$$

We could get equations of motion of a system in two ways, one from the Lagrangian (the Euler-Lagrange equations)

$$\partial_t \partial_{\dot{q}} L = \partial_q L$$

or from the Hamiltonian

$$\partial_q H = -\dot{p} \qquad \partial_p H = \dot{q}$$

Both Lagrangian and Hamiltonian, like potential energy, are simple functions and so independent of the coordinate system chosen. But the Hamiltonian, being the total energy, is not Lorentz-invariant: Lorentz invariance has the form

$$E^2 - p^2 = m^2$$

which does not change under a Lorentz transformation between energy and momentum (just as $x^2 + y^2 = r^2$ does not change under a rotation).

The Lagrangian, while not necessarily Lorentz invariant (e.g., in the nonrelativistic simple harmonic oscillator $L = (mv^2 + kx^2)/2$) it can be, and should be to describe a relativistic system. In fact, once we have a Lorentz-invariant Lagrangian, anything derived from it will obey special relativity.

So we will work with Lagrangians and with generalized coordinates. With the latter we can derive a Hamiltonian, which is useful when we want to discuss energy in some particular coordinate frame.

Our field equation so far gives us only one "coordinate", ϕ. This is also going to be an operator, because it will be formed from the Fourier transform of creation and annihilation operators. It will need a complementary "coordinate", π, to satisfy the Lagrangian/Hamiltonian formalisms. We will also find commutation relationships between ϕ and π, analogous to those between X and P in Note 18.

But moving from discrete coordinates such as X and P or even such as the generalized coordinates q and p, to fields which have values throughout timespace, will require us to elaborate the Lagrangian formalism, especially the Euler-Lagrange equations. We are going from function $L(q, p)$ of coordinates to functions $\mathcal{L}(\phi, \pi)$ of fields which extend throughout timespace or energentum space, and will have to be taken as *densties*, per unit volume.

So we change notation from L (and H) (which I've used only temporarily in this Note) to $\mathcal{L}$ (and $\mathcal{H}$) (which I used in Book 8c but not there meaning densities).

And we must elaborate the Euler-Lagrange equations. The new ones can be derived by a variational principle, as Feynman does for L, cited but not discussed in Book 8c. I'll just give the results and then work an example for familiarization.

The conjugate momentum is

$$\pi = \partial_{\partial_t \phi} \mathcal{L}$$

It is only part of the extended Euler-Lagrange

$$\partial_\mu (\partial_{\partial_\mu \phi} \mathcal{L}) = \partial_\phi \mathcal{L}$$

with repeated index summed $\mu = 0, 1, 2, 3$.

Let's see how it works. Here is a Lagrangian

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}(\not{\partial}\phi \ \not{\partial}\phi - m^2\phi^2)$$

(Physicists sometimes pull Lagrangians out of the air, like this one, with the intention of arriving at a known equation, or they may derive a Lagrangian systematically.)

Let's see that this gives the Klein-Gordon equation. Expand the Lagrangian

$$\frac{1}{2}(\not{\partial}\phi \ \not{\partial}\phi - m^2\phi^2) = \frac{1}{2}((\partial_t\phi)^2 - (\partial_x\phi)^2 - (\partial_y\phi)^2 - (\partial_z\phi)^2 - m^2\phi^2)$$

Expand each side of the Euler-Lagrange equation.

$$\begin{aligned} \partial_\mu \partial_{\partial_\mu \phi} \mathcal{L} &= \partial_t(\partial_{\partial_t \phi}\mathcal{L}) + \partial_x(\partial_{\partial_x \phi}\mathcal{L}) + \partial_y(\partial_{\partial_y \phi}\mathcal{L}) + \partial_z(\partial_{\partial_z \phi}\mathcal{L}) \\ &= \partial_t\partial_t\phi - \partial_x\partial_x\phi - \partial_y\partial_y\phi - \partial_z\partial_z\phi \\ &= \not{\partial}\phi \ \not{\partial}\phi \\ \partial_\phi \mathcal{L} &= -m^2\phi \end{aligned}$$

Equate the two sides

$$(\not{\partial}\phi \ \not{\partial} + m^2)\phi = 0$$

which is the Klein-Gordon equation.

The conjugate momentum is

$$\pi = \partial_{\partial_t \phi} \mathcal{L} = \partial_t \phi$$

just the time slope of ϕ.

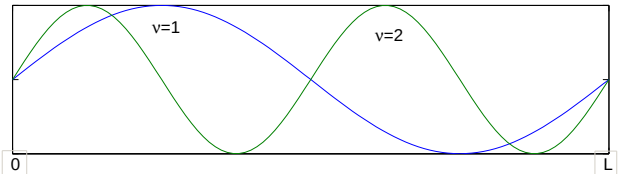
To find the commutator $[\phi, \pi]_-$ we must construct ϕ, and hence π, from its momentum states.

$$\phi_\ell = \frac{1}{\sqrt{L^3}} \sum_k \frac{1}{\sqrt{2\omega_k}} (e^{i\not{k} \cdot \not{\ell}\, 2\pi/L} U_{(-)k} + e^{-i\not{k} \cdot \not{\ell}\, 2\pi/L} D_{(-)k})$$

Here I've discretized the wavenumbers

$$k_x = \frac{2\pi}{L} \nu_x$$

etc., for integers ν_x, ν_y, ν_z, with L the length of the side of the containing volume of space in each dimension. This way ϕ_k is zero on any side of the L^3 box:



And I've discretized the spatial positions so we are calculating $\phi_{\ell_x \ell_y \ell_z}$ instead of a continuous $\phi(x, y, z)$. That way we can use the discrete Fourier transform (DFT) of Week 9, Notes 1, 2 and especially Excursion *FT with vector* k, ℓ. Note that $\not{k}$ and $\not{\ell}$ are three-dimensional, not four, and represent the spatial components. The time component is fixed to one instant so ω plays the role given k_x in the last paragraph of that Excursion. In particular, we can multiply the Fourier coefficients by any function of ω, and for reasons which will become clear we choose the inverse square root.

We just found out that

$$\pi = \partial_t \phi = \frac{i}{\sqrt{L^3}} \sum_k \sqrt{\frac{\omega_k}{2}} (-e^{i\not{k} \cdot \not{\ell}\, 2\pi/L} U_{k(-)} + e^{-i\not{k} \cdot \not{\ell}\, 2\pi/L} D_{k(-)})$$

The $D_{k(-)}$ and $U_{k(-)}$ are just ladder operators, occupying different parts of Foch space for different values of k. Thus we have the commutators

$$\begin{aligned} [D_{(-)k}, D_{(-)k'}]_- &= 0 = [U_{(-)k}, U_{(-)k'}]_- \\ [D_{(-)k}, U_{(-)k'}]_- &= \delta_{kk'} \end{aligned}$$

Then

$$\begin{aligned} [\phi_\ell, \pi_{\ell'}]_- &= \phi_\ell \pi_{\ell'} - \pi_{\ell'} \phi_\ell \\ &= \frac{i}{2L^3} \sum_{k,k'} \sqrt{\frac{\omega_{k'}}{\omega_k}} (e^{i(k' \cdot \ell' + k \cdot \ell) 2\pi/d} [U_{(-)k'}, U_{(-)k}]_- \\ &\quad + e^{i(k' \cdot \ell' - k \cdot \ell) 2\pi/d} [U_{(-)k'}, D_{(-)k}]_- \\ &\quad + e^{i(k \cdot \ell - k' \cdot \ell') 2\pi/d} [U_{(-)k}, D_{(-)k'}]_- \\ &\quad + e^{-i(k' \cdot \ell' + k \cdot \ell) 2\pi/d} [D_{(-)k'}, D_{(-)k}]_-) \\ &= \frac{i}{2L^3} \sum_k \left(e^{ik \cdot (\ell' - \ell) 2\pi/d} + e^{-ik \cdot (\ell' - \ell) 2\pi/d} \right) \\ &= \frac{i}{L^3} \delta_{\ell, \ell'} \end{aligned}$$

In the four lines following the second equality, the first and last vanish because of the commutators and the middle two require that $k' = k$ because those commutators equal $\delta_{k,k'}$. Thus the ω_k and $\omega_{k'}$ cancel out in the third equality, justifying our use of $1/\sqrt{\omega_k}$ at the beginning.

The final result comes from the property of the Fourier transform that we developed in the Notes 1 and 2 of Week 9.

This result is essentially the commutator of position and momentum that we saw in Note 18 on the quantum harmonic oscillator. It is missing the factor of $\hbar$ but that can be put back in after fizzmezh analysis. And it has an inverse factor of L^3. This is essentially the volume of the space (once we take the argument over to continuous variables), which is appropriate because we've called both ϕ and π probability *densities*, that is to say, per unit volume.

Similar arguments show that $[\phi_\ell, \phi_{\ell'}]_- = 0 = [\pi_\ell, \pi_{\ell'}]_-$, completing the position-momentum commutation relations. (Instead of there being a sum of two equal terms when $\ell = \ell'$ there is a difference.)

25. The Yukawa potential. What new physics can we extract from the field theory we have so far? How about the force between two objects? Yukawa in 1935 worked out the effective potential due to the exchange of a field excitation (a particle) between two nucleii: the excitation is created at one nucleon and annihilated at the other.

We need some preliminary work with the Klein-Gordon equation. First the *static* case, $\partial_t \phi = 0$

$$(\not{\partial} \cdot \not{\partial} + m^2)\phi = 0 \rightarrow (-\not{\nabla} \cdot \not{\nabla} + m^2)\phi = 0$$

Now $\not{\nabla} \cdot \not{\nabla} = \partial_x^2 + \partial_y^2 + \partial_z^2 = \nabla^2$, the last being the conventional symbol. For a spherical system we can write ∇^2 in two ways, the first being the usual form in spherical coordinates

$$\begin{aligned} \nabla^2 \phi &= \frac{1}{r^2} \partial_r (r^2 \partial_r \phi) \\ &= \frac{1}{r} \partial_r^2 (r\phi) \end{aligned}$$

which we can show to be equivalent to each other via

$$\nabla^2 = \frac{2}{r}\partial_r + \partial_r^2$$

Then the static Klein-Gordon equation

$$(\nabla^2 - m^2)\phi = 0$$

is

$$\partial_r^2 (r\phi) = m^2 (r\phi)$$

and has the solution

$$\phi(r) = \frac{e^{-mr}}{r}$$

This is the form taken by the Yukawa potential, but we must see why it is a *potential*.

A more general solution to the Klein-Gordon equation makes use of a formal "inverse", called a "Green's function": the $\delta(x)\delta(y)\delta(z)$ plays the role of the identity matrix in this continuous function.

$$(-\nabla^2 + m^2) G(\vec{x}) = \delta(x)\delta(y)\delta(z)$$

We will need the Fourier transform of $G(\vec{x})$

$$G(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\not{k} \cdot \not{x}} \; G(\vec{k})$$

(See Excursion *Continuous FT* in Week 9 for the conversion from discrete to continuous Fourier transforms.)

Since

$$\begin{aligned}\bigtriangledown^2 e^{i\vec{k}\cdot\vec{x}} &= (\partial_x^2+\partial_y^2+\partial_z^2)e^{i(k_x x+k_y y+k_z z)}\\ &= -(k_x^2+k_y^2+k_z^2)e^{i(k_x x+k_y y+k_z z)}\end{aligned}$$

putting the two equations for G together gives

$$(k^2+m^2)G(\vec{x})=1$$

(the converse Fourier transform has been used) so

$$G(\vec{x})=\int\frac{d^3k}{(2\pi)^3}\frac{e^{i\vec{k}\cdot\vec{x}}}{k^2+m^2}$$

I'm now going to be sloppy and equate these two results, with the constant of integration being the additional divisor, 4π. This is to avoid getting into calculating contour integrals in the 2D number plane. (It is done properly in Note 28.) So

$$\int\frac{d^3k}{(2\pi)^3}\frac{e^{i\vec{k}\cdot\vec{x}}}{k^2+m^2}=\frac{e^{-mr}}{4\pi r}$$

Now let's look at what a field excitation between two nucleii does to the energy of the field. Because we are calculating *energy* differences we'll need to use the Hamltonian. But we'll need only the energetically small *interaction Hamiltonian*, $\mathcal{H}_I$, giving the change in the field as it interacts with the two nucleii, not the original Hamiltonian, $\mathcal{H}_0$, of the field alone.

The next Note describes the "perturbation" we'll now use for the slightly perturbed Hamiltonian $\mathcal{H}_0+\epsilon\mathcal{H}_I$. Here we just plunge in because only one result from the next Note is needed and the discussion here will motivate the discussion there of perturbation.

Since the unperturbed field, $\phi(x)$, contains creation and annihilation operators for the excitations, it must form part of the interaction Hamiltonian, together with an indication of the positions of the two nucleii, and a "coupling constant" g which gives the strength of the interaction and, of course, the right fizzmezh to make $\mathcal{H}_I$ an energy. We'll use delta functions to locate the nucleii at x_1 and x_2.

$$\begin{aligned}\epsilon\mathcal{H}_I &= g(\delta(x-x_1)+\delta(x-x_2))\phi(\vec{x})\\ &= \frac{g}{\sqrt{V}}\sum_k\frac{1}{\sqrt{2\omega_k}}(U_{(-)k}(e^{i\vec{k}\cdot\vec{x}_1}+e^{i\vec{k}\cdot\vec{x}_2})+D_{(-)k}(e^{-i\vec{k}\cdot\vec{x}_1}+e^{-i\vec{k}\cdot\vec{x}_2}))\end{aligned}$$

(The ϵ is just a very small number, used in the next Note, which may be considered part of g.)

Now we must think of $\mathcal{H}_I$ (and $\epsilon\mathcal{H}_I$) as a matrix of elements linking the states $|\,..,n_k,..>$ to each other. Remember that the one term, n_k, represents an infinite vector representing the number of excitations in the kth quantum harmonic oscillator by a 1 in the n_kth position.

$D_{(-)k}$ and $U_{(-)k}$ operate only on this one term, with the effect (Note 18) that

$$<..,n_k+1,..\,|\,\epsilon\mathcal{H}_I\,|\,..,n_k,..>=\frac{g}{\sqrt{V}}\sqrt{\frac{n_k+1}{2\omega_k}}(e^{-i\vec{k}\cdot\vec{x}_1}+e^{-i\vec{k}\cdot\vec{x}_2})$$

and

$$<..,n_k,..\,|\,\epsilon\mathcal{H}_I\,|\,..,n_k+1,..>=\frac{g}{\sqrt{V}}\sqrt{\frac{n_k+1}{2\omega_k}}(e^{i\vec{k}\cdot\vec{x}_1}+e^{i\vec{k}\cdot\vec{x}_2})$$

(for each different position, k, in the bras and kets, of course) and all the other matrix elements are zero.

Before and after its excitation by the nucleii the field will be in its *ground state* $\mid 0 >= \mid 0, 0, .. >$. So $n_k + 1 = 1$ for each k.

The next Note will consider first and second order perturbations of the field, giving energy changes

$$\Delta E =< 0 \mid \epsilon \mathcal{H}_I \mid 0 >$$

to first order—and you can see from the above that this must be 0—and, to second order,

$$\Delta E = \sum_k \frac{< 0 \mid \epsilon^2 \mathcal{H}_I \mid 0, .., n_k, 0, .. >< 0, .., n_k, 0, .. \mid \epsilon^2 \mathcal{H}_I \mid 0 >}{-\omega_k}$$

where $n_k = 1$ for each k and the sum is over all positions k in the Fock vector $\mid 0, .., n_k, 0, .. >$. That is,

$$\begin{aligned}
\Delta E &= < 0 \mid \epsilon^2 \mathcal{H}_I \mid 1, 0, 0, .. >< 1, 0, 0, .. \mid \epsilon^2 \mathcal{H}_I \mid 0 > /(-\omega_0) \\
&\quad + < 0 \mid \epsilon^2 \mathcal{H}_I \mid 0, 1, 0, .. >< 0, 1, 0, .. \mid \epsilon^2 \mathcal{H}_I \mid 0 > /(-\omega_1) \\
&\quad + < 0 \mid \epsilon^2 \mathcal{H}_I \mid 0, 0, 1, .. >< 0, 0, 1, .. \mid \epsilon^2 \mathcal{H}_I \mid 0 > /(-\omega_2) \\
&\quad + \cdots \\
&= \frac{-g^2}{V} \sum_k \frac{1}{2\omega_k^2} (e^{i\vec{k} \cdot (\vec{x}_1 - \vec{x}_1)} + e^{i\vec{k} \cdot (\vec{x}_2 - \vec{x}_2)} + e^{i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)} + e^{i\vec{k} \cdot (\vec{x}_2 - \vec{x}_1)})
\end{aligned}$$

Since we are interested for now in the interaction *between* the nucleii at x_1 and x_2 we'll ignore the "self-energy" terms involving $x_1 - x_1$ and $x_2 - x_2$. (These caused serious problems historically, since the sums they produce are infinite, requiring "renormalization", but we'll pass them by.)

So we have two terms left, involving the separation $\Delta x = x_2 - x_1$ between the two nucleii.

$$\begin{aligned}
\Delta E &= \frac{-g^2}{2V} \sum_k \frac{1}{\omega_k^2} (e^{i\vec{k} \cdot \Delta\vec{x}} + e^{-i\vec{k} \cdot \Delta\vec{x}}) \\
&= \frac{-g^2}{V} \sum_k \frac{1}{k^2 + m^2} e^{i\vec{k} \cdot \Delta\vec{x}} \\
&\to \frac{-g^2}{(2\pi)^3} \int d^3k \frac{e^{i\vec{k} \cdot \Delta\vec{x}}}{k^2 + m^2} \\
&= -g^2 \frac{e^{-mr}}{4\pi r}
\end{aligned}$$

where line 2 notes that the sums over k and $-k$ are both the same and that $\omega_k^2 = k^2 + m^2$; line 3 converts the discrete to the continuous Fourier transform (Excursion *Continuous FT* of Week 9); and line 4 uses the result earlier in this Note, with $r^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$ being the distance in space between the two nucleii.

The two terms involving $x_1 - x_2$ and $x_2 - x_1$ can be interpreted respectively as the creation of an excitation at nucleus 2 and its annihilation at nucleus 1, and vice-versa.

This energy difference is the *Yukawa potential*—it generalizes the classical Coulomb and gravitational $1/r$ potential to "force carriers"—the field excitations—of nonzero mass m. Thus it has a short range, dictated by the size of m, as opposed to the infinite range of the $1/r$ forces—which have carriers of 0 mass, the photon and the "graviton". This is the beginnings of an explanation

of the short ranges of the nuclear forces. From the field point of view, all forces are mediated by particles, massless in the case of the long-range forces, massy in the case of short-range forces.

The Yukawa potential gives rise to an attractive force because of the minus sign. What we have worked out here pertains to an excitation of "spin 0": it is described by a scalar which is symmetric under *any* rotation of the axes. Field theory can also go on to show that a spin-1 excitation must be repulsive between like charges—e.g., the electromagnetic force—and a spin-2 excitation—e.g., the graviton—must be attractive again.

Next, however, we will move to spin-1/2. After first looking into perturbations.

26. Perturbation approximations. The negative sign that makes the Yukwa potential attractive comes directly from the second-order perturbation. So we must examine the perturbation approximation at least that far.

Suppose we have some matrix, $H_)$, whose eigenvalues, $E_n^{(0)}, n = 1, 2, ..,$ and eigenvectors, ϕ_n, we know. (I'm using notation which suggests Hamiltonians, energies and wavefunctions, but H_0 can be any hermitian matrix, say a 2-by-2 shear matrix.)

Now suppose that a small matrix, ϵH_I, is added to H_0 (with the smallness captured by the very small number ϵ), and that we want to find the eigenvalues, E_n, of the sum $H = H_0 + \epsilon H_I$.

We expand E_n in a series of powers of ϵ, starting with the original eigenvalues $E_n^{(0)}$.

$$E_n = E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + ..$$

We'll need the new eigenvectors, ψ_n. Since the original eigenvectors, ϕ_n, give an orthonormal basis for the space of eigenvectors, we can write ψ_n as a linear combination of the ϕ_n

$$\psi_n = \sum_k c'_{nk} \phi_k$$

Since we expect ψ_n to be close to ϕ_n we'll rewrite this as

$$\psi_n = \phi_n + \sum_{k \neq n} c_{nk} \phi_k$$

(or just omit the $k \neq n$ and restrict $c_{nn} = 0$). Actually, for the ψ_n to be normalized we must multiply the sum on the right by a normalizing constant. But this will cancel out on both sides of the equations to follow, and anyway can be shown to be 1 to within order ϵ^2.

We expand the coefficients c_{nk} also in powers of ϵ

$$c_{nk} = \epsilon c_{nk}^{(1)} + \epsilon^2 c_{nk}^{(2)} + ..$$

(Why is there no $c_{nk}^{(0)}$?)

So

$$H\psi_n = E_n \psi_n$$

expands into products of infinite series

$$(H_0 + \epsilon H_I)(\phi_n + \sum_{k \neq n} c_{nk}\phi_k) = (E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + ..)(\phi_n + \sum_{k \neq n} c_{nk}\phi_k)$$

where we know

$$H_0 \phi_n = E_n^{(0)} \phi_n$$

and

$$\phi_j \cdot \phi_k = \delta_{kj}$$

By equating on both sides the coefficients of each power of ϵ, we can work our way up the chain to whatever power of ϵ we wish. We'll go to ϵ^2: we'll need $E_n^{(1)}$ and for $E_n^{(2)}$ we'll need $c_{nk}^{(1)}$.

First, ϵ^0. Equating the infinite series gives

$$H_0\phi_n = E_n^{(0)}\phi_n$$

which we already knew.

Next, ϵ^1.

$$H_I\phi_n + H_0\sum_{k\neq n} c_{nk}^{(1)}\phi_k = E_n^{(1)}\phi_n + E_n^{(0)}\sum_{k\neq n} c_{nk}^{(1)}\phi_k$$

Note that the c_{nk} (to all orders) are just coefficients multiplying vectors ϕ_k. They do not form a matrix which operates on ϕ_k the way, say, H_0 does. So we can take H_0 into the sum, e.g.,

$$H_0\sum_{k\neq n} c_{nk}^{(1)}\phi_k = \sum_{k\neq n} c_{nk}^{(1)}H_0\phi_k = \sum_{k\neq n} c_{nk}^{(1)}E_k^{(0)}\phi_k$$

We can pick out the $E_n^{(1)}$ and the $c_{nk}^{(1)}$ by using the dot product with ϕ_n and $\phi_j, j\neq n$, respectively.

$$\phi_n\cdot H_I\phi_n + \phi_n\cdot\sum_{k\neq n} c_{nk}^{(1)}E_k^{(0)}\phi_k = E_n^{(1)} + \phi_n\cdot\sum_{k\neq n} c_{nk}^{(1)}E_n^{(0)}\phi_k$$

so

$$E_n^{(1)} = \phi_n\cdot H_I\phi_n$$

$$\phi_j\cdot H_I\phi_n + \phi_j\cdot\sum_{k\neq n} c_{nk}^{(1)}E_k^{(0)}\phi_k = E_n^{(1)}\phi_j\cdot\phi_n + \phi_j\cdot\sum_{k\neq n} c_{nk}^{(1)}E_n^{(0)}\phi_k$$

so

$$\phi_j\cdot H_I\phi_n + c_{nj}^{(1)}E_j^{(0)} = c_{nj}^{(1)}E_n^{(0)}$$

or

$$c_{nj}^{(1)} = \frac{\phi_j\cdot H_I\phi_n}{E_n^{(0)} - E_j^{(0)}}$$

Finally, ϵ^2.

$$H_0\sum_{k\neq n} c_{nk}^{(2)}\phi_k + H_I\sum_{k\neq n} c_{nk}^{(1)}\phi_k = E_n^{(1)}\sum_{k\neq n} c_{nk}^{(1)}\phi_k + E_n^{(2)}\phi_n$$

Taking the dot product with ϕ_n removes the sums after H_0 and $E_n^{(1)}$ leaving

$$\begin{aligned}
E_n^{(2)} &= \phi_n\cdot E_n^{(2)}\phi_n \\
&= \phi_n\cdot H_I\sum_{k\neq n} c_{nk}^{(1)}\phi_k \\
&= \sum_{k\neq n}(\phi_n\cdot H_I\phi_k)c_{nk}^{(1)} \\
&= \sum_{k\neq n}\frac{(\phi_n\cdot H_I\phi_k)(\phi_k\cdot H_I\phi_n)}{E_n^{(0)} - E_k^{(0)}}
\end{aligned}$$

This is the result we used in Note 25 with $E_n^{(0)} = 0$ and $E_k^{(0)} = \omega_k$ giving the $-\omega_k$ denominator and the numerator as the sum of products of dot products, which we wrote, e.g.,

$$\phi_0\cdot H_I\phi_k \rightarrow < 0 \mid H_I \mid 0,..,n_k,0,.. >$$

There, the resulting energy, to second order, was

$$E_0^{(0)} + \epsilon^2 E_0^{(2)}$$

27. Fermions. The Klein-Gordon equation can be factored

$$\begin{aligned} 0 &= -(\not{\partial}\cdot\ \not{\partial} + m^2)\psi \\ &= (i\ \not{\partial} - m)(i\ \not{\partial} + m)\psi \end{aligned}$$

(or $\not{\partial}\cdot\ \not{\partial} + m^2 = (\not{\partial} - im)(\not{\partial} + im)$ but the above is the convention).
Let's see what happens with (the Dirac equation)

$$(i\ \not{\partial} - m)\psi = 0$$

First, it is related to the other factor

$$\begin{aligned} 0 = ((i\ \not{\partial} - m)\psi)^\dagger &= \psi^\dagger(i\overleftarrow{\not{\partial}} - m)^\dagger \\ &= \overline{\psi}d_0(id_0\overleftarrow{\partial_0} - id_j\overleftarrow{\partial_j} - m)^\dagger \\ &= \overline{\psi}d_0(-id_0^\dagger\overleftarrow{\partial_0} + id_j^\dagger\overleftarrow{\partial_j} - m) \end{aligned}$$

where the $\overleftarrow{\partial}$ indicates that the slope must be applied to the left, allowing us to reverse the order of the terms when we took the transpose; and where it is useful to define

$$\overline{\psi} \stackrel{\text{def}}{=} \psi^\dagger d_0$$

Multiplying on the right by d_0 (and recalling $d_0^2 = 1, d_0 d_j = -d_j d_0$ for $j = 1, 2, 3$, and $d_0^\dagger = d_0$, while $d_j^\dagger = -d_j$), we continue

$$\begin{aligned} 0 &= \overline{\psi}d_0(-id_0^\dagger\overleftarrow{\partial_0} + id_j^\dagger\overleftarrow{\partial_j} - m)d_0 \\ &= -\overline{\psi}(id_0^\dagger\overleftarrow{\partial_0} - id_j^\dagger\overleftarrow{\partial_j} + m) \\ &= -\overline{\psi}(i\overleftarrow{\not{\partial}} + m) \end{aligned}$$

The Lagrangian

$$\mathcal{L} = \overline{\psi}(i\ \not{\partial} - m)\psi$$

gives us back the Dirac equation using Euler-Lagrange with respect to ψ

$$0 = \partial_\mu(\partial_{\partial_\mu\overline{\psi}}\mathcal{L}) = \partial_{\overline{\psi}}\mathcal{L} = (i\ \not{\partial} - m)\psi$$

since the 2D-number slope (see the next Note) $\partial_{x+iy}(x - iy) = 0$ and $\overline{\psi} = \psi^\dagger d_0$ with $\psi^\dagger$ transposing ψ and changing the sign of i in it: so $\partial_{\overline{\psi}}\psi = 0$ and $\partial_\mu(\partial_{\partial_\mu\overline{\psi}}\mathcal{L}) = 0$ above.

The "momentum" corresponding to ψ is essentially $\psi^\dagger$:

$$\begin{aligned} \pi &= \partial_{\partial_0\psi}\mathcal{L} \\ &= \partial_{\partial_0\psi}\psi^\dagger d_0(id_0\partial_0 - id_j\partial_j - m)\psi \\ &= i\psi^\dagger \end{aligned}$$

But now we'll find that it is not the commutator $[\psi, \pi]_-$ which is significant but the anticommutator

$$[\psi, \pi]_+$$

We'll see this in the context of the Fourier transform of these functions with *anti*commuting creation and annihilation operators as coefficients: the fields now describe fermions.

The argument parallels that for the simple field in Note 24 but is complicated by the fact that the Dirac equation involves 4-by-4 matrices $d_0, d_j, j = 1, 2, 3$, and of course the identity matrix I in what must now be mI. So there will be four solutions, each a 4-element vector.

(That 4D timespace gives rise to 4-by-4 matrices and 4-element vectors is something of a coincidence: a 5D "timespace" would also give rise to 4-by-4s, etc., just as both the 2D and 3D space give rise to the 2-by-2 Pauli matrices.)

If we introduce plane waves and let

$$\psi(x) = \begin{cases} u_1(p)e^{-ik\!\!\!/\cdot x\!\!\!/} \\ u_2(p)e^{-ik\!\!\!/\cdot x\!\!\!/} \\ v_1(p)e^{ik\!\!\!/\cdot x\!\!\!/} \\ v_2(p)e^{ik\!\!\!/\cdot x\!\!\!/} \end{cases}$$

then the Dirac equation becomes, for $s = 1, 2$ and $j = 1, 2, 3$ and repeated js summed

$$\begin{aligned} 0 &= (i\ \partial\!\!\!/ - m)u_s e^{-k\!\!\!/\cdot x\!\!\!/} \\ &= (id_0\partial_0 + id_j\partial_j - m)u_s e^{-ik_0x_0 + ik_jx_j} \\ &= (k_0d_0 - k_jd_j - m)u_s \\ &= (k\!\!\!/ - m)u_s \\ 0 &= (i\ \partial\!\!\!/ - m)v_s e^{k\!\!\!/\cdot x\!\!\!/} \\ &= (id_0\partial_0 + id_j\partial_j - m)v_s e^{ik_0x_0 - ik_jx_j} \\ &= (-k_0d_0 + k_jd_j - m)v_s \\ &= -(k\!\!\!/ + m)v_s \end{aligned}$$

So the Fourier transform has four pieces, involving u_1, u_2, v_1 and v_2, and there are four Fourier transforms, one for each component of the vectors.

$$\psi_\alpha(x) = \frac{1}{\sqrt{V}}\sum_k \frac{1}{\sqrt{2\omega_k}}((D_{(+)1k}(u_{1k})_\alpha + D_{(+)2k}(u_{2k})_\alpha)e^{-ik\!\!\!/\cdot x\!\!\!/} + (U_{(+)1k}(v_{1k})_\alpha + U_{(+)2k}(v_{2k})_\alpha)e^{ik\!\!\!/\cdot x\!\!\!/})$$

and similarly for the four components of the conjugate momenta

$$\psi^\dagger_\alpha(x) = \frac{1}{\sqrt{V}}\sum_{k'} \frac{1}{\sqrt{2\omega_{k'}}}((D^\dagger{}_{(+)1k'}(u_{1k'})^*_\alpha + D^\dagger{}_{(+)2k'}(u_{2k'})^*_\alpha)e^{ik\!\!\!/'\cdot x\!\!\!/} + (U^\dagger{}_{(+)1k'}(v_{1k'})^*_\alpha + U^\dagger{}_{(+)2k'}(v_{2k'})^*_\alpha)e^{-ik\!\!\!/'\cdot x\!\!\!/})$$

where the * on a number reverses the sign of i in that number (2D-conjugate or complex conjugate) and † on a matrix is the usual Hermitian transpose.

The $D_{(+)}$ and $U_{(+)}$ are the ladder operators—annihilation and creation, respectively—and, as we found in Note 20, we can have two of each for 4-by-4 matrices, just what we need here.

The Hermitian conjugates

$$D^\dagger{}_{(+)} = U_{(+)}$$

so, of course,

$$U^\dagger{}_{(+)} = D_{(+)}$$

and they obey the anticommutation relationships

$$\begin{aligned} [D_{(+)sk}, D_{(+)tk'}]_+ &= 0 \\ [U_{(+)sk}, U_{(+)tk'}]_+ &= 0 \\ [D_{(+)sk}, U_{(+)tk'}]_+ &= \delta_{st}\delta_{kk'} \end{aligned}$$

Here the δ_{st} express the anticommutation relationships from Note 20. The $\delta_{kk'}$ comes from the valid assumption that we can extend the Fock space for anticommutators indefinitely so that different values of k are independent of each other in the same sense.

With this background, we can show, summing over repeated α,

$$[\psi_\alpha(x,t),\psi^\dagger_\alpha(x',t)]_+ = \frac{1}{V}\delta_{xx'}$$

i.e.,

$$[\psi_\alpha(x,t),\pi_\alpha(x',t)]_+ = \frac{i}{V}\delta_{xx'}$$

the generalized coordinates and momenta for fermions anticommute in exact analogy with the commutators of the same for bosons.

The anticommutator $[\psi_\alpha,\psi^\dagger_\alpha]_+$ above is a sum

$$\frac{1}{2V}\sum_{kk'}\frac{1}{\sqrt{\omega_k\omega_{k'}}}(\cdots)$$

of four components

$$\begin{array}{ccccccccccc}
[(D_{(+)1k}(& u_{1k})_\alpha + & D_{(+)2k}(& u_{2k})_\alpha) & e^{-ik\not{x}}, & (D^\dagger{}_{(+)1k'}(& u_{1k'})^*_\alpha + & D^\dagger{}_{(+)2k'}(& u_{2k'})^*_\alpha) & e^{ik\not{x}}]_+ \\
[(D & u & D & u &)e^- & (U^\dagger & v^* & U^\dagger & v^* &)e^-]_+ \\
[(U & v & U & v &)e^+ & (D^\dagger & u^* & D^\dagger & u^* &)e^+]_+ \\
[(U & v & U & v &)e^+ & (U^\dagger & v^* & U^\dagger & v^* &)e^-]_+
\end{array}$$

where I've left the subscripts off the last three lines to make it easy to spot the differences: the $1k, 2k, 1k', 2k'$ and α are identical to those in the corresponding positions of the first line. Similarly for the exponents apart from signs.

Now, because of the transposes and commutators, the middle two lines vanish and the outer two lines produce $\delta_{kk'}$ which reduce the double sum $\sum_{kk'}$ to a single sum $\sum_k$ containing exponentials $e^{\pm ik(x-x')}$. These sum to 0 except when $x' = x$ and everything works out as in Note 24, provided the "spinors" are normalized

$$\begin{array}{rcl}
(u_{rk})^\dagger(u_{sk})^\dagger = & 2\omega_k\delta_{rs} & = (v_{rk})^\dagger(v_{sk})^\dagger \\
(v_{rk})^\dagger(u_{sk})^\dagger = & 0 & = (u_{rk})^\dagger(v_{sk})^\dagger
\end{array}$$

28. Slopes and antislopes of 2D numbers, etc. Note 27 said that $\partial_{x+iy}x - iy = 0$ and Note 25 mentioned contour integrals. This Note backtracks to look at these ideas.

Considering functions of 2D numbers, it would make sense to say that a slope exists only if a function is a *continuous* transformation of the number. That is because slopes are approximated by ratios of differences involving the functional values at some point and a nearby point, with "nearby" taken to be arbitrarily close.

A rotation, $\cos\theta + i\sin\theta = e^{i\theta}$, or a rotation-plus-expansion, $a(\cos\theta + i\sin\theta) = ae^{i\theta}$. are obvious examples of continuous transformations: θ can be made arbitrarily small. A reflection is not continuous.

The 2D conjugate operation

$$x + iy \to x - iy$$

is a reflection in the x=axis. Other reflections can be written easily. In the y-axis

$$x + iy \to -x + iy;$$

in the $x = y$ line

$$x + iy \to y + ix.$$

For reflections, especially the conjugate operation, we would not expect a slope to be defined. This says it better than our earlier assertion that the slope is 0.

We can harden up the above discussion by saying that the "slope" of a 2D function of 2D numbers is defined only if the slope gives the same number (which can itself be 2D) for *every* direction of "nearby". This is going to turn out to be a very restrictive condition mathematically, but one satisfied by many functions which describe physical things, and so very useful for physics.

So if we write

$$\begin{aligned} z &= x + iy \\ f(z) &= u(x,y) + iv(x,y) \end{aligned}$$

and so

$$\begin{aligned} \Delta z &= \Delta x + i\Delta y \\ \Delta f &= \Delta u + i\Delta v \end{aligned}$$

and if we insist that $\Delta f/\Delta z$ is independent of the direction of Δz, then we can look at two examples, (i)

$$\begin{aligned} \Delta y &= 0 \\ \frac{\Delta f}{\Delta z} &= \frac{\Delta u}{\Delta x} + i\frac{\Delta v}{\Delta x} \end{aligned}$$

i.e.,

$$\partial_z f = \partial_x u + i\partial_x v$$

and (ii)

$$\begin{aligned} \Delta x &= 0 \\ \frac{\Delta f}{\Delta z} &= \frac{\Delta u}{i\Delta y} + i\frac{\Delta v}{i\Delta y} \end{aligned}$$

i.e.,

$$\partial_z f = -i\partial_y u + \partial_y v$$

And we're asking these to be the same

$$\begin{aligned} \partial_y v &= \partial_x u \\ \partial_y u &= -\partial_x v \end{aligned}$$

Let's look at the relationship between Δz and Δf in matrix form (the "Jacobian" matrix).

$$\begin{aligned} \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix} &= \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &= \begin{pmatrix} \partial_x u & -\partial_x v \\ \partial_x v & \partial_x u \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &= \begin{pmatrix} \partial_y v & \partial_y u \\ -\partial_y u & \partial_y v \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \end{aligned}$$

Both of these equivalent matrices have the form of rotation-plus-expansion.

So generalized rotations are the *only* legal functional mapping, given our constraint on direction independence. (I *said* it was restrictive.)

The constraints are known as the *Cauchy-Riemann* equations. Functions obeying these constraints are called "analytical" functions. They are very special. They lead, by the way, to the result that both $u(x, y)$ and $v(x, y)$ must be solutions to the 2D Laplace equation

$$\begin{aligned} \bigtriangledown^2 u &= \partial_x \partial_x u + \partial_y \partial_y u = 0 \\ \bigtriangledown^2 v &= \partial_x \partial_x v + \partial_y \partial_y v = 0 \end{aligned}$$

What kind of functions satisfy these constraints? Well, $z = x + iy$ does for a start. And then multiplying it by $z = x + iy$, to give z^2, is effectively to rotate-and-expand it, so z^2 satisfies the constraints. So does z^3, etc.

Thus, any series

$$a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots$$

will satisfy the Cauchy-Riemann equations, and hence so will any function expressible as such a series. This includes exponential functiona e^z and e^{iz}, and so the trigonometric functions $\cos(z)$ and $\sin(z)$, etc., and hyperbolic functions $\cosh(z)$ and $\sin(z)$, etc. And so on.

An exception must be made for the logarithmic functions $\ln(z)$, $\log(z)$, $\lg(z)$, etc.: these have $1/z$-like poles and those poles introduce interesting new considerations. I'll skip over the full development of "functions of complex variables" but we will encounter $1/z$-type poles and their usefulness very soon.

If the slopes of analytical functions don't care about directions, it is plausible that their antislopes don't either.

We're going to have to switch to more conventional notation and vocabulary because we're working in two dimensions. Ordinarily in 1D

$$\text{antislope}_x f(x) \approx \sum f(x_i) \Delta x$$

the area under $f(x)$ over the range of x_i summmed, with $\Delta x = x_{i+1} - x_i$, the distance between successive x_i; or, with specific limits,

$$\text{antislope}_{x=a:b} f(x) \approx \sum_{x_0=a}^{x_n=b} f(x_i) \Delta x$$

The conventional notation reflects this

$$\int f(x) dx = \text{antislope}_x f(x)$$

or

$$\int_a^b f(x) dx = \text{antislope}_{x=a:b} f(x)$$

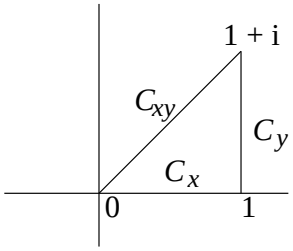
The conventional vocabulary calls this an *integral*. dx has the same meaning a Δx except shrunk to infinitesimally small.

In 2D we must specify a curve, C, in the x-y (or z) plane, along which the function, $f(z)$ is "integrated".

$$\int_C f(z) dz$$

in general, different curves give different results.

Let's integrate $x - iy$ over two curves



$$
\begin{aligned}
C_x &: \quad y = 0, dy = 0 \\
&\int_{c_x} (x - iy)(dx + idy) = \int_0^1 x dx = \frac{x^2}{2} \mid_0^1 = \frac{1}{2} \\
C_y &: \quad x = 1, dx = 0 \\
&\int_{c_y} (x - iy)(dx + idy) = \int_0^1 (1 - iy) i dy = iy + \frac{y^2}{2} \mid_0^1 = i + \frac{1}{2} \\
C_{xy} &: \quad x = y, dx = dy \\
&\int_{c_{xy}} (x - iy)(dx + idy) = \int x dx (1 - i)(1 + i) = 2\frac{x^2}{2} \mid_0^1 = 1
\end{aligned}
$$

So going from 0 to 1+i along C_x then C_y gives $1 + i$ while going straight along C_{xy} gives 1.

On the other hand, the integral of $x + iy$ does not depend on the paths: both ways give i as the result.

What are the general conditions for the integral to be independent of the path?

$$\int_C (u + iv)(dx + idy) = \int_C (udx - vdy) + i \int_C (udy + vdx)$$

If both of the integrals are exactly the step of some other function $\phi(x, y)$, as in

$$\Delta\phi = \partial_x \phi \Delta x + \partial_y \phi \Delta y$$

then we have, for the first component integral

$$u\Delta x - v\Delta y = \partial_x \phi \Delta x + \partial_y \phi \Delta y$$

so

$$
\begin{aligned}
u &= \partial_x \phi \\
v &= -\partial_y \phi
\end{aligned}
$$

and therefore

$$\partial_y u = \partial_y \partial_x \phi = \partial_x \partial_y \phi = -\partial_x v$$

For the second component integral we can say a similar thing about some other function ϕ'

$$v\Delta x + u\Delta y = \partial_x \phi' \Delta x + \partial_y \phi' \Delta y$$

so

$$
\begin{aligned}
v &= \partial_x \phi' \\
u &= \partial_y \phi'
\end{aligned}
$$

and

$$\partial_x u = \partial_x \partial_y \phi' = \partial_y \partial_x \phi' = \partial_y v$$

But these two results

$$\begin{aligned} \partial_y v &= \partial_x u \\ \partial_y u &= -\partial_x v \end{aligned}$$

are just the Cauchy-Riemann equations: the condition that slope direction makes no difference is just the condition that the path of integration makes no difference.

For an analytic function $f(y)$ and *any* curve C from z_1 to z_2

$$\int_C f(z)dz = \text{antislope}_{z=z_2} f - \text{antislope}_{z=z_1} f$$

and there is a simple relationship between 2D integrals and antislopes.

What's really remarkable, then, is that for any *closed* curve, any analytical function integrates to zero.

$$\oint f(z)dz = 0$$

(I *said* analytical functions were restrictive.)

But if the closed curve surrounds a *pole* of $f(z)$ then $f(z)$ is not analytical *everywhere inside* and the integral is not zero.

Let's try a curve which is the unit circle with centre at the origin, and the function z^n (which we found above is analytical for any non-negative n).

We write $z = re^{-\theta}$ in polar coordinates.

$$z^n dz = r^n e^{in\theta}(e^{i\theta}dr + ire^{i\theta}d\theta)$$

and on the unit circle $r = 1$, $dr = 0$, so

$$\begin{aligned} \oint z^n dz &= \int_0^{2\pi} d\theta i e^{i(n+1)\theta} \\ &= \begin{cases} \int_0^{2\pi} i d\theta = 2\pi i & n = -1 \\ \frac{e^{i(n+1)\theta}}{n+1} \,|_0^{2\pi} = 0 & \text{otherwise} \end{cases} \end{aligned}$$

where we've considered all integer values of n: positive, 0 or negative.

So we've seen another remarkable thing:

$$\oint \frac{dz}{z} = 2\pi i$$

and this will be true for any closed curve surrounding the origin. Somehow the pole at the origin affects the integral of $1/z$ on any closed curve around it.

But not just any pole: higher-order poles such as $1/z^2, 1/z^3$, etc., have no effect on integrals around these poles, which are all zero.

The pole need not be at the origin. We can shift variables.

$$\begin{aligned} z' &= z - a \qquad dz' = dz \\ 2\pi i &= \oint \tfrac{dz}{z} = \oint \tfrac{dz'}{z'} = \oint \tfrac{dz}{z-a} \end{aligned}$$

so now the pole is at some other point, a. The result is still true as long as the integration path surrounds a.

Note that $1/z$ is the slope of the logarithm ln z (in 1D): these first-order poles link to the special nature, mentioned above, of the logarithm.

This is not all. We can extend these results to any function with a first-order pole. Let's consider an analytic function $f(z)$ and a curve which is a circle of small radius ϵ centred at point a: $z = a+\epsilon e^{i\theta}$, $dz = i\epsilon e^{i\theta} d\theta$, so

$$\begin{aligned} \oint \frac{f(z)}{z-a} dz &= \int_0^{2\pi} d\theta f(a+\epsilon e^{i\theta}) \frac{i\epsilon e^{i\theta}}{\epsilon e^{i\theta}} \\ &= \int_0^{2\pi} i d\theta f(a+\epsilon e^{i\theta}) \\ \epsilon \stackrel{\longrightarrow}{\rightarrow} 0 & \int_0^{2\pi} d\theta i f(a) \\ &= 2\pi i f(a) \end{aligned}$$

Thus, if we integrate around point a, a function with a simple pole at a, the answer is $2\pi i$ times the function less (i.e., multiplied by) the $z-a$ part, evaluated at $z = a$.

We can use this to evaluate the "contour integral" which we did sloppily in Note 25. First consider

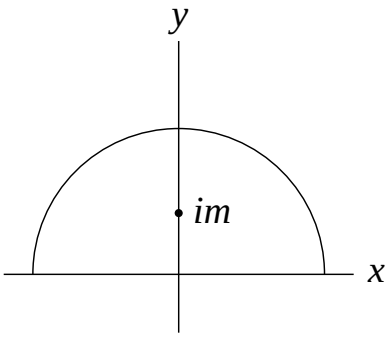
$$\oint dz \frac{ze^{izr}}{z^2+m^2} = \oint dz \frac{ze^{izr}}{(z+im)(z-im)}$$

with the closed curve surrounding the simple pole $z = im$ in the upper half of the z plane.

Now we know that the integral is the value at $z = im$ of

$$2\pi i \frac{ze^{izr}}{(z+im)} = 2\pi i \frac{ime^{-mr}}{2im} = \frac{2\pi i}{2} e^{-mr}$$

We can make this curve a "D" along the horizontal axis and with a semicircular roof.



Now if we let the "D" become infinitely large, the contour integral becomes the sum of two parts,

$$\int_{-\infty}^{\infty} dz \frac{ze^{izr}}{z^2+m^2}$$

and the same integral on a semicircle of infinite radius. But on this semicircle $z = x + iy$ with x and y both arbitrarily large, so

$$e^{izr} = e^{i(x+iy)r} = e^{ixr} e^{-yr}$$

becomes infinitely small and we're left with

$$\int_{-\infty}^{\infty} dz \frac{ze^{izr}}{z^2+m^2} = \frac{2\pi i}{2} e^{-mr}$$

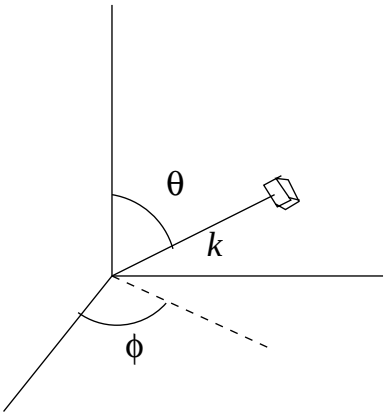
Now comes some ordinary work. We must just relate

$$\frac{1}{(2\pi)^3} \int d^3k \frac{e^{i\vec{k}\cdot\vec{x}}}{k^2+m^2}$$

to the integral we've just solved, and we will have cleaned up the sloppiness of Note 25.

First we switch to polar coordinates in k

$$d^3k = dk k d\theta k \sin\theta d\phi$$



Then, since we're integrating $\vec{k}$ in all directions and sizes, we can chose the direction of $\vec{x}$ in $\vec{x}\cdot\vec{k}$ to be anything we like, so we choose $\vec{x}$ to lie in the z direction, in which case $\vec{x}\cdot\vec{k} = kr\cos\theta$ for $r^2 = x^2+y^2+z^2$. Note, for the next step, that if $q = \cos\theta$ then $dq = -\sin\theta d\theta$, q going from 1 to -1 as θ goes from 0 to π.

So, here goes.

$$\begin{aligned}
\frac{1}{(2\pi)^3}\int d^3k \frac{e^{ikr\cos\theta}}{k^2+m^2} &= \frac{1}{(2\pi)^3}\int_0^\infty dk \int_0^{2\pi} d\theta \sin\theta\, 2\pi k^2 \frac{e^{ikr\cos\theta}}{k^2+m^2} \\
&= \frac{1}{(2\pi)^2}\int_0^\infty dk \frac{k^2}{k^2+m^2}\int_1^{-1}(-dq)e^{ikrq} \\
&= \frac{1}{(2\pi)^2}\int_0^\infty dk \frac{k^2}{k^2+m^2}\frac{1}{ikr}e^{ikrq}\,|_{-1}^{1} \\
&= \frac{1}{(2\pi)^2 ir}\int_0^\infty dk \frac{k}{k^2+m^2}(e^{ikr}-e^{-ikr}) \\
&= \frac{1}{(2\pi)^2 r}\int_0^\infty dk \frac{k}{k^2+m^2} 2\sin(kr) \\
&= \frac{1}{2\pi r}\int_{-\infty}^\infty dk \frac{k}{k^2+m^2}\frac{\sin(kr)}{2\pi} \\
&= \frac{1}{2\pi r}\int_{-\infty}^\infty dk \frac{k}{k^2+m^2}\mathrm{Re}\left(\frac{e^{ikr}}{2\pi i}\right) \\
&= \frac{1}{2\pi r}\mathrm{Re}\left(\frac{1}{2\pi i}\int_{-\infty}^\infty dk \frac{ke^{ikr}}{k^2+m^2}\right)
\end{aligned}$$

Here each step should be self-explanatory, apart from the jiggling around from $e^{ikr} - e^{-ikr}$ to $\sin(kr)$ to the "real" (x-axis) part of e^{ikr}/i; note that we could change the integration limit from $0 : \infty$ to $-\infty : \infty$, getting rid of a factor 2, because $k \sin(kr)$ is an even function.

We can now change variables from k to z and use the contour integral, giving

$$\frac{1}{2\pi r} \mathrm{Re}\left(\frac{1}{2\pi i}\frac{2\pi i}{2}\right) = \frac{1}{4\pi r} e^{-mr}$$

This is the result I claimed in Note 25: the Yukawa potential.

29. Charge conservation and antimatter. We return to bosons and commutators and the simple field of Note 24, which we extend to two fields, ϕ_1 and ϕ_2. The Lagrangian is just the sum

$$\mathcal{L} = \frac{1}{2}(\not{\partial}\phi_1 \ \not{\partial}\phi_1 - m^2\phi_1^2) + \frac{1}{2}(\not{\partial}\phi_2 \ \not{\partial}\phi_2 - m^2\phi_1^2)$$

What makes two fields interesting is that they can combine into one 2D-number field.

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$$

for which the Lagrangian is

$$\mathcal{L} = \not{\partial}\phi^\dagger \ \not{\partial}\phi - m^2\phi^\dagger\phi$$

(Check that these two forms of $\mathcal{L}$ are the same.)

There are now two sets of annihilation and creation operators, $D_{(-)1k}, U_{(-)1k}, D_{(-)2k}$ and $U_{(-)2k}$ (these are much easier to construct than the multiple $D_{(+)}, U_{(+)}$ for fermions) with $U_{(-)jk} = D_{(-)jk}^\dagger$ and which commute or not as usual

$$\begin{aligned}\left[D_{(-)jk}, U_{(-)j'k'}\right]_- &= \delta_{jj'}\delta_{kk'} \\ \left[D_{(-)jk}, D_{(-)j'k'}\right]_- &= 0 = \left[U_{(-)jk}, U_{(-)j'k'}\right]_-\end{aligned}$$

They can be combined

$$\begin{aligned} D_{(-)k} &= \tfrac{1}{\sqrt{2}}(D_{(-)1k} + iD_{(-)2k}) \quad U_{(-)k} = D_{(-)k}^\dagger = \tfrac{1}{\sqrt{2}}(U_{(-)1k} - iU_{(-)2k}) \\ \hat{D}_{(-)k} &= \tfrac{1}{\sqrt{2}}(D_{(-)1k} - iD_{(-)2k}) \quad \hat{U}_{(-)k} = \hat{D}_{(-)k}^\dagger = \tfrac{1}{\sqrt{2}}(U_{(-)1k} + iU_{(-)2k}) \end{aligned}$$

and you can show

$$[D_{(-)k}, U_{(-)k'}]_- = \delta_{kk'} = \left[\hat{D}_{(-)k}, \hat{U}_{(-)k'}\right]_-$$

and all the other commutators vanish.

The relationship between ϕ and the ladder operators is by Fourier transform, as for the simple boson field in Note 24 or the fermion field in Note 27.

$$\phi(x) = \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}}(D_{(-)k} e^{-ik\not{\cdot}\not{x}} + \hat{U}_{(-)k} e^{-ik\not{\cdot}\not{x}})$$

from which

$$\phi^\dagger(x) = \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}}(\hat{D}_{(-)k} e^{-ik\not{\cdot}\not{x}} + U_{(-)k} e^{-ik\not{\cdot}\not{x}})$$

and we'll see soon why we've mixed the hatted operators with the others.

The benefit of putting these two fields into 2-number form is that we can see a symmetry of the Lagrangian. If we alter ϕ to $\phi + \Delta\phi = \phi e^{-iq\theta}$ and hence $\phi^\dagger$ to $\phi^\dagger + \Delta\phi^\dagger = \phi^\dagger e^{iq\theta}$, for some constant q and any angle θ, the Lagrangian does not change, $\Delta\mathcal{L} = 0$.

This might remind you of the electromagnetic Schrödinger equation which we found and simulated in Notes 13 and 14.

In Note 38 of Book 8c (Part IV) we saw that symmetries of the Lagrangian imply conservation laws and conserved quantities. These were simple applications of Noether's therem for regular Lagrangians, Now we are dealing with *field* Lagrangian densities, and the corresponding Euler-Lagrange equations of Note 24:

$$\begin{aligned} \partial_\mu(\partial_{\partial_\mu\phi}\mathcal{L}) &= \partial_\phi\mathcal{L} \\ \partial_\mu(\partial_{\partial_\mu\phi^\dagger}\mathcal{L}) &= \partial_{\phi^\dagger}\mathcal{L} \end{aligned}$$

which we must now apply to both fields. (They will, of course, give two Klein-Gordon equations, one for each field in Note 24.)

We want to show that something is conserved—and what that something is—because $\Delta\mathcal{L} = 0$ when a transformation such as $\phi \to \phi + \Delta\phi$ causes the Lagrangian $\mathcal{L} \to \mathcal{L} + \Delta\mathcal{L}$.

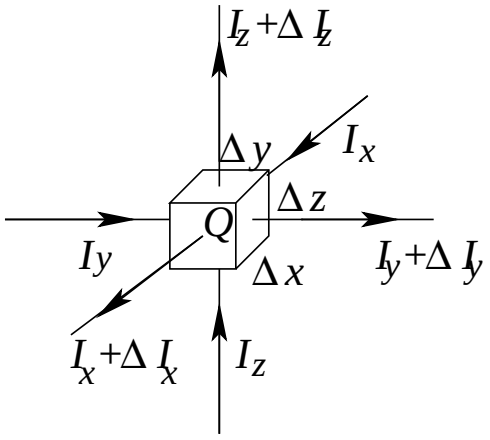
We can look for something else to become 0, and relate that something to $\Delta\mathcal{L}$. An appropriate "something else" is the 4-divergence of a current density ("density", i.e., per unit volume, or per unit area, because the Lagrangians we're using in field theory are also densities).

For a 4-current density J_μ the 4-divergence is

$$0 = \not\partial \cdot \not J = \partial_t J_t + \partial_x J_x + \partial_y J_y + \partial_z J_z$$

where $J_t = \rho$ the density per unit volume of the charge (fizzmezh Q/L^3) and J_x, J_y, J_z are components of the current density per unit area flowing away from the charge (fizzmezh $I/L^2 = Q/TL^2$).

We can see this relationship by looking at a small box in a substance with a charge density ρ and currrent density (J_x, J_y, J_z).



(The figure shows charge $Q = \rho\Delta x\Delta y\Delta z$ rather than charge density ρ, and current, $I_x = J_x\Delta y\Delta z$, etc., rather than current density, J_x, etc.)

To sketch the argument, suppose $\Delta I_x = 0 = \Delta I_y$ so only the z-current has any effect on the net charge inside the box. This charge, Q, diminishes because of the net outgoing current (we suppose ΔI_z is positive)

$$-\partial_t\rho\Delta x\Delta y\Delta z = -\partial_t Q = \Delta I_z = \Delta J_z\Delta x\Delta y$$

so

$$\partial_t \rho \Delta z + \Delta J_z = 0$$

i.e.,

$$\partial_t \rho + \partial_z J_z = 0$$

For currents in all directions

$$\partial_t J_t + \partial_x J_x + \partial_y J_y + \partial_z J_z = 0$$

So we need $\Delta\mathcal{L} \propto \not{\partial}\cdot\ \not{J}$ for some 4-current density J_μ. Then the conserved quantity will be the charge $Q = \int d^3x \partial_t J_t$.

We need the expansion of $\Delta\mathcal{L}(\phi, \phi^\dagger, \partial_\mu\phi, \partial_\mu\phi^\dagger)$:

$$\Delta\mathcal{L} = (\partial_\phi\mathcal{L})\Delta\phi + (\partial_{\phi^\dagger}\mathcal{L})\Delta\phi^\dagger + (\partial_{\partial_\mu\phi}\mathcal{L})\Delta\partial_\mu\phi + (\partial_{\partial_\mu\phi^\dagger}\mathcal{L})\Delta\partial_\mu\phi^\dagger$$

For this we need

$$\begin{aligned}
\Delta\phi &= -iq\theta\phi \\
\Delta\phi^\dagger &= iq\theta\phi^\dagger \\
\Delta\partial_\mu\phi &= -iq\theta\partial_\mu\phi \\
\Delta\partial_\mu\phi^\dagger &= iq\theta\partial_\mu\phi^\dagger
\end{aligned}$$

for infinitesimal ϕ so $e^{-iq\theta} \approx 1 - iq\theta$, $e^{iq\theta} \approx 1 + iq\theta$ and θ is an internal "angle" which does not depend on t, x, y or z so $\partial_\mu\theta = 0$.

Finally we will try out

$$J_\mu = -iq((\partial_{\partial_\mu\phi}\mathcal{L}) - (\partial_{\partial_\mu\phi^\dagger}\mathcal{L}))$$

Some algebra, using the Euler-Lagrange equations, expands

$$\begin{aligned}
\not{\partial}\cdot\ \not{J} &= -iq((\partial_\mu\partial_{\partial_\mu\phi}\mathcal{L})\phi + (\partial_{\partial_\mu\phi}\mathcal{L})(\partial_\mu\phi) - (\partial_\mu\partial_{\partial_\mu\phi^\dagger}\mathcal{L})\phi^\dagger - (\partial_{\partial_\mu\phi^\dagger}\mathcal{L})(\partial_\mu\phi^\dagger)) \\
&= -iq((\partial_\phi\mathcal{L})\phi - (\partial_{\phi^\dagger}\mathcal{L})\phi^\dagger + (\partial_{\partial_\mu\phi}\mathcal{L})(\partial_\mu\phi) - (\partial_{\partial_\mu\phi^\dagger}\mathcal{L})(\partial_\mu\phi^\dagger))
\end{aligned}$$

Going back to $\Delta\mathcal{L}$ and the approximations for $\Delta\phi$, etc.,

$$\begin{aligned}
\Delta\mathcal{L} &= -iq\theta((\partial_\phi\mathcal{L})\phi - (\partial_{\phi^\dagger}\mathcal{L})\phi^\dagger + (\partial_{\partial_\mu\phi}\mathcal{L})(\partial_\mu\phi) - (\partial_{\partial_\mu\phi^\dagger}\mathcal{L})(\partial_\mu\phi^\dagger)) \\
&= \theta\ \not{\partial}\cdot\ \not{J}
\end{aligned}$$

So $\Delta\mathcal{L} = 0$ (from the symmetry under the $e^{-iq\theta}$ transformation) implies $\not{\partial}\cdot\ \not{J} = 0$ and J_μ is the conserved 4-current.

The above argument is a step closer to the fully general Noether theorem from the discussion in Book 8c, Note 38. Here we have used quantum field theory rather than simple one-particle quantum mechanics.

What does it mean? Now we'll see why we mixed $D_{(-)k}$ with $\hat{U}_{(-)k}$ and $\hat{D}_{(-)k}$ with $U_{(-)k}$ in expanding $\phi(x)$ and $\phi^\dagger(x)$.

First let's write the 4-current density for our 2D field Lagrangian

$$J_\mu = -iq((\partial_\mu\phi^\dagger)\phi - (\partial_\mu\phi)\phi^\dagger)$$

These are the two terms we must expand as ladder operators in the expression

$$\begin{aligned}
Q &= \int d^3x J_t = -iq \int d^3x((\partial_t\phi^\dagger)\phi - (\partial_t\phi)\phi^\dagger) \\
i(\partial_\mu\phi^\dagger)\phi) &= \frac{1}{V}\sum_k \sqrt{\frac{\omega_k}{4\omega_k}}(D_{(-)}e^- - \hat{U}_{(-)}e^+)(\hat{D}_{(-)}e^- + U_{(-)}e^+) \\
-i(\partial_\mu\phi)\phi^\dagger) &= \frac{1}{V}\sum_k \sqrt{\frac{\omega_k}{4\omega_k}}(U_{(-)}e^+ - \hat{D}_{(-)}e^-)(D_{(-)}e^+ + \hat{U}_{(-)}e^-)
\end{aligned}$$

I've left out some repeated subscripts and exponentials, and I've used $k_0 = \omega_k$—so that the ω_k factors will cancel out.

Multiplying these out and adding them together we have commutators multiplying exponentials $e^{\pm 2i\not{k}\cdot\not{x}}$ which go to zero:

$$[D_{(-)}, \hat{D}_{(-)}]_- = 0 = [\hat{U}_{(-)}, U_{(-)}]_-$$

And we have exponentials which cancel to 1 multiplying nonzero anticommutators:

$$DU + UD = 2UD + 1$$

and

$$\hat{D}\hat{U} + \hat{U}\hat{D} = 2\hat{U}\hat{D} + 1$$

and these are subtracted from one another so the 1s cancel too. The 2s cancel the $1/\sqrt{4}$ and the integral over all space cancels the $1/V$, so the final answer, putting back the q we left out temporarily, is

$$Q = q\sum_k (U_{(-)k} D_{(-)k} - \hat{U}_{(-)k}\hat{D}_{(-)k})$$

From Note 18 we know that $U_{(-)}D_{(-)}$ are just the counting operator N, so

$$Q = q(N - \hat{N})$$

where N counts one type of particle and $\hat{N}$ counts a second type of particle.

The conserved quantity Q is the difference in total charges: N particles of the first type, each with charge q; $\hat{N}$ particles of the second type, each with charge $-q$.

Dirac eventually interpreted these two types of particle as matter and antimatter.

The original expansion of $\phi(x)$ in terms of $D_{(-)k}$ and $\hat{U}_{(-)k}$ can be interpreted as removing a particle of charge q or adding an antiparticle of charge $-q$: it reduces the overall charge. On the other hand, $\phi^\dagger(x)$ increases the overall charge by removing a $-q$ or adding a q.

30. Relativistic quantum field theory redux, so far. Moving from single particle quantum mechanics to relativistic quantum field theory using annihilation and creation operators has introduced us to important differences between bosons and fermions, to forces mediated by carrier particles including short-range nuclear forces with massive carriers and hints about whether forces are attractive or repulsive depending on the spin of the boson carriers, and to antimatter and charge conservation.

Quantum field theory offers great benefits. But it is also very hard to calculate with—we've already seen some gentle hints about the length of calculations.

This exploration has taken us through the physics of the 1930s. Then there was a 10-year hiatus (which included World War II and the Manhattan Project) during which the computational

difficulties led many physicists to contemplate abandoning quantum field theory.

Part V. Functional Integrals

31. Path amplitudes.
32. Functionals.
33. Gaussian integrals.
34. Diagrams and QED.
35. Chirality and electroweak.
36. Green's functions.
37. Propagators.
38. Quantum Computing.
39. Binary Fourier transform.
40. Quantum Fourier transform.
41. Finding periods.
42. Quantum key distribution.
43. No cloning.
44. Database search.
45. Detecting and correcting errors.
46. Nonlocality: Einstein-Podolsky-Rosen.
47. Building a quantum computer.

II. **The Excursions**

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

1. Note 18 discusses bosons and fermions. Refer back to Week 5 and to the Excursion in Week 7a that discusses Pauli's discovery of the effect of special relativity on quantum phase.

2. **Heisenberg uncertainty.**Jordan [Jor86, pp.129–32] uses matrices such as P and X of Note 18 to derive Heisenberg's uncertainty principle from the commutator, namely if $[X,P]_- = i\hbar$ then $\sqrt{< (X- < X >)^2 >}\sqrt{< (P- < P >)^2 >} \geq \hbar/2$. This follows directly from $< A^2 >< B^2 > \geq |\ (1/2) < [A,B]_- >|^2$ which he shows for "real" matrices A and B. You will recognize the standard deviations of X and P, which give the uncertainties: see below.
   In Jordan's terms, a matrix represents a quantity, all possible values of which that can result from a measurement of that quantity are in its set of eigenvalues. If all eigenvalues are real (which they are for Hermitian matrices) that matrix is "real". If they are all non-negative, the matrix is non-negative, and so on.
   In chapter 13 Jordan gives rules for mean values, $< \cdots >$, which are sums of the eigenvalues weighted by probabilities: linearity $< aA + bB >= a < A > +b < B >$ where A and B are matrices and a and b are numbers; sign $<$ negative matrix $> =$ negative number; commutation for real B and D, $[B,D]_-$ is imaginary and hence $< BD >=< DB >^*$; for commuting B and D, $(B+iD)(B-iD) = B^2 + D^2 \geq 0$; and for real matrices and any complex number w, $(K+wL)(K-w^*L) \geq 0$. You should check this last one out: it is used in the proof of Heisenberg.

3. **More uncertainty.** With the 2-by-2 reflection matrices representing spin ($\text{spin}_j = (1/2)f_j$ for $j = x, y, z$) we can be more concrete than with the infinite matrices of the previous Excursion. The eigenvalues of f_j are 1 and -1 in each case, so let's say the norm of a matrix $|\ f_j\ | = \max\ |$ eigenvalues $| = 1$

We can compute the expected values explicitly for an arbitrary state $(c, se^{-i\alpha})$ (c could be multiplied by a phase factor, too, but the net effect would be an overall phase factor which will make no difference).

$$\begin{aligned}
< f_z > &= (c, se^{-i\alpha}) \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \begin{pmatrix} c \\ se^{\alpha} \end{pmatrix} = c^2 - s^2 = c_2 \\
< f_x > &= (c, se^{-i\alpha}) \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} c \\ se^{\alpha} \end{pmatrix} = cs(e^{\alpha} + e^{-i\alpha}) = 2cs\cos\alpha = s_2\cos\alpha \\
< f_y > &= (c, se^{-i\alpha}) \begin{pmatrix} & -i \\ i & \end{pmatrix} \begin{pmatrix} c \\ se^{\alpha} \end{pmatrix} = -ics(e^{\alpha} - e^{-i\alpha}) = -2csi^2\sin\alpha = s_2\sin\alpha
\end{aligned}$$

and we note that the sum of squares of the expectation values $< f_z >^2 + < f_x >^2 + < f_y >^2 = c_2^2 + s_2^2(\cos^2\alpha + \sin^2\alpha) = 1$. This says that if $c_2 = \pm 1$, say, so $s_2 = 0$ and the state (c, s) (never mind α) is parallel to the z-axis, then (twice) the spin is certainly ± 1 in the z-direction and completely uncertain ($< f_x >= 0 =< f_y >$) in the other perpendicular directions.

The derivation of the uncertainty principle in the previous Excursion does not depend on the size of the matrices so we can check out the commutator of, say, f_z with some reflection intermediate between f_z and f_x, say.

$$\left[\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix}\right]_- = s_2 \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} = is_2 \begin{pmatrix} & -i \\ i & \end{pmatrix}$$

and the norm of this is s_2 (the eigenvalues are ± 1). Thus the uncertainty is 0 if both matrices are the same but a maximum if the matrices are orthogonal to each other in the sense that the two reflections described are at 45 degrees.

For 2-by-2 matrices with eigenvalues ± 1 the expectation values and the associated probabilities have a 1-to-1 relationship. Let p_{-1} be the probability of measuring -1 and p_1 be the probability of measuring 1.

$$\begin{aligned}
p_1 + p_{-1} &= 1 \\
p_1 - p_{-1} &= < f_j >
\end{aligned} \qquad \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_{-1} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1- \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ < f_j > \end{pmatrix} = \begin{pmatrix} 1 \\ < f_j > \end{pmatrix} \qquad \begin{pmatrix} p_1 \\ p_{-1} \end{pmatrix}$$

For quantum physics, more fundamental than probabilities are *amplitudes* and we can get these from the components of the state. Working in 2D with the intermediate matrix and state (c', s') we transform both to axes in which the matrix is diagonalized.

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} c' \\ s' \end{pmatrix} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \begin{pmatrix} cc' + ss' \\ cs' - sc' \end{pmatrix} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \begin{pmatrix} c_- \\ s_- \end{pmatrix}$$

So the amplitude for this state giving the value 1 on measurement is c_- which is the cosine of the difference between the angle (c, s) of the matrix and the angle (c', s') of the state. The amplitude for finding -1 is s_-. The corresponding probabilities are the squares of the amplitudes.

Alternatively, we can project (c', s') onto (c, s) using $P = (1 + F)/2$ where F is a reflection and P is the corresponding projection.

$$\begin{aligned}
\frac{1}{2}\left(I + \begin{pmatrix} c_2 & s_2 \\ s_2 & -c_2 \end{pmatrix}\right) &= \frac{1}{2} \begin{pmatrix} 1 + c^2 - s^2 & 2cs \\ 2cs & 1 - c^2 + s^2 \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} 2c^2 & 2cs \\ 2cs & 2s^2 \end{pmatrix}
\end{aligned}$$

$$
\begin{aligned}
& & = & \begin{pmatrix} c \\ s \end{pmatrix} (c, s) \\
\begin{pmatrix} c \\ s \end{pmatrix} (c, s) \begin{pmatrix} c' \\ s' \end{pmatrix} & = & \begin{pmatrix} c \\ s \end{pmatrix} (cc' + ss') \\
& = & c_- \begin{pmatrix} c \\ s \end{pmatrix}
\end{aligned}
$$

This gives the amplitude c_-. Similarly the complementary, orthogonal projection $P = (1 - F)/2$ gives amplitude s_- for projection on the eigenvector with value -1. Projection in this way corresponds better to the measurement process in which we measure one particular outcome (e.g., putting light through a filter which polarizes it in one particular direction), irreversibly losing information about the original state of the system.

4. Write out the multiplication table for the three base reflections f_1, f_2 and f_3 of Note 19 and show that it has $8 = 2^3$ elements, including 1, f_{12} and f_{123}.

5. Persuade yourself that f_3 of Note 19 is a reflection in 3D, mapping, say, $(x, y, z) \to (x, -z, -y)$:

$$
\begin{pmatrix} & -i \\ i & \end{pmatrix} \begin{pmatrix} x + iy \\ z \end{pmatrix} = \begin{pmatrix} -iz \\ -y + ix \end{pmatrix}
$$

6. **Sphere of reflections.** a) Show that $f = xf_x + yf_y + zf_z$ (Note 19) is a reflection if $x^2 + y^2 + z^2 = 1$, i.e., its square is the identity, $f^2 = I$. That is, the three basic reflections generate a sphere of radius $r = \sqrt{x^2 + y^2 + z^2} = 1$, every point of the surface of which also represents a reflection.
b) Since a point on the surface is a reflection, a point in the interior of the sphere is a *mix* of reflections: consider a disc in the plane perpendicular to the radius from (0,0,0) to the interior point (x, y, z) and bounded by the sphere of radius 1. Any two points at opposite ends of a diameter of the disc are on the sphere and hence (pure) reflections, say f_a and f_b. The interior point at the centre of the disc is thus the "mixed" reflection $(f_a + f_b)/2$.
In fact, *any* disc containing the interior point in question can lead to a pair of pure reflections which mix, $af_a + bf_b$, $a + b = 1$, to give it. Indeed, any triple of points also mix, $af_a + bf_b + cf_c$, $a + b + c = 1$; or any number of points on the boundary of such a disc.
c) Since $p = (I + f)/2$ is a projection for any reflection f, the sphere of reflections is also a sphere of (pure) projections, and the interior points represent mixes of projections. (Show that $p^2 = p$.)
d) The projection onto a normalized vector (c, s) is given by the matrix

$$
\begin{pmatrix} c \\ s \end{pmatrix} (c\ s) = \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix}
$$

so a mixed projection is given by, say,

$$
a \begin{pmatrix} c_1 \\ s_1 \end{pmatrix} (c_1\ s_1) + b \begin{pmatrix} c_2 \\ s_2 \end{pmatrix} (c_2\ s_2)
$$

with $a + b = 1$. (Note the difference: $a + b = 1$ but $c_j^2 + s_j^2 = 1$.)
Either of these is called a *density* matrix: the pure projection is a *pure* density matrix, corresponding to, say, a deterministic "probability distribution"; the mixed projection is a *mixed* density matrix, corresponding to the more familiar probability distribution a, b with $a + b = 1$ and neither one zero.
e) Show that when (c_2, s_2) is orthogonal to (c_1, s_1) (e.g., $(c_2, s_2) = (s_1, -c_1)$), these vectors are the eigenvectors, and a and b the eigenvalues, of the (symmetrical) density matrix. This

applies also if $a+b$ is not restricted to unity and the resulting matrix is any symmetrical 2-by-2 matrix.
f) For the non-orthogonal case, explore

$$a\begin{pmatrix}1\\0\end{pmatrix}(1\ 0)+\frac{b}{2}\begin{pmatrix}1\\1\end{pmatrix}(1\ 1)$$

which has eigenvalues $(a+b\pm\sqrt{a^2+b^2})/2$ and eigenvectors (x,y) satisfying $0=(ax+by\mp x\sqrt{a^2+b^2})/2$; for $a=3/5$ and $b=4/5$ check that the original matrix equals

$$\frac{6}{25}\begin{pmatrix}2\\1\end{pmatrix}(2\ 1)+\frac{1}{25}\begin{pmatrix}1\\-2\end{pmatrix}(1\ -2)$$

(Note that I've made $a^2+b^2=1$ just to get rid of square roots: this example is not a density matrix.)

7. **Density matrices and inverse tensor product.** The tensor product of two density matrices is a density matrix and can be decomposed into the two original density matrices by taking restricted traces.

$$\begin{pmatrix}a&c\\b&d\end{pmatrix}\overleftarrow{\times}\begin{pmatrix}a'&c'\\b'&d'\end{pmatrix}=\left(\begin{array}{cc|cc}aa'&ac'&ca'&cc'\\ab'&ad'&cb'&cd'\\\hline ba'&bc'&da'&dc'\\bb'&bd'&db'&dd'\end{array}\right)$$

Since the 2-by-2 matrices are density matrices, their traces $a+d=1$ and $a'+d'=1$; it follows that the trace of the 4-by-4 is also 1 and so the tensor product is also a density matrix.
The decomposing operators are the "left trace" $\mathrm{Tr_L}$ and the "right trace" $\mathrm{Tr_R}$ which combine the elements of the 4-by-4 as follows.

$\mathrm{Tr_L}$

1 3	
1 3	
2 4	
2 4	

$\mathrm{Tr_R}$

1 3	
2 4	
	1 3
	2 4

Show that these give back the left and right operands of $\overleftarrow{\times}$ respectively.

8. **Certainty and entanglement.** In Note 5 of Book 9c (Part I) we introduced density matrices as a way of representing probability distributions: such a density matrix is diagonal, with the diagonal elements being the probabilities, and summing, of course, to 1. In the case of certainty, one of the probabilities is 1 and the rest 0, and the density matrix is a projection. Generally, for non-diagonal density matrices, a density matrix which is a projection represents certainty in the following, quantum-mechanical, sense. A quantum-mechanical state, a vector (c,s) with $c^2+s^2=1$, gives a result with certainty when measured in the direction given by c and s. This is because (c,s) is the eigenvector of the density matrix

$$\begin{pmatrix}c^2&cs\\cs&s^2\end{pmatrix}\begin{pmatrix}c\\s\end{pmatrix}=\left(\begin{pmatrix}c\\s\end{pmatrix}(c,s)\right)\begin{pmatrix}c\\s\end{pmatrix}=\begin{pmatrix}c\\s\end{pmatrix}\left((c,s)\begin{pmatrix}c\\s\end{pmatrix}\right)=\begin{pmatrix}c\\s\end{pmatrix}(c^2+s^2)$$

and the density matrix, a projection, represents the measurement (think of a polarizing filter). For classical probabilities (diagonal density matrices) a 4-by-4 projection will decompose (using the restricted traces of the previous Excursion) into two projections: certainty decomposes

to certainty. For example

$$\begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$$

In general, however, a 4-by-4 which is a projection need not decompose into projections. Here is a "Bell state"

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \times \frac{1}{\sqrt{2}}(1,0,0,1) = \frac{1}{2} \begin{pmatrix} 1 & & & 1 \\ & & & \\ & & & \\ 1 & & & 1 \end{pmatrix}$$

which *is* a projection (it must be, but you can square it to check) but decomposes (using the templates of the previous Excursion) into

$$\frac{1}{2} \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \quad (\text{using } \mathrm{Tr_L}) \quad \text{and} \quad \frac{1}{2} \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \quad (\text{using } \mathrm{Tr_R})$$

Neither of these is a projection. Each can be made up of a mix of two projections.

$$\frac{1}{2} \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}$$

There are four Bell states like this (without the $1/\sqrt{2}$): $(1,0) \overleftarrow{\times} (1,0) \pm (0,1) \overleftarrow{\times} (0,1)$ and $(1,0) \overleftarrow{\times} (0,1) \pm (0,1) \overleftarrow{\times} (1,0)$.
A state such as one of these Bell states is called "entangled": two particles together form a state but cannot be separated into two definite states, one for each particle.
Examples of entangled states are: two electrons of opposite spins created from some spin-0 system (note that the $-$ sign makes the state antisymmetrical so that swapping the electrons changes the sign)

$$\frac{1}{\sqrt{2}}((1,0) \overleftarrow{\times} (0,1) - (0,1) \overleftarrow{\times} (1,0));$$

or two photons polarized in the same direction (symmetrical)

$$\frac{1}{\sqrt{2}}((1,0) \overleftarrow{\times} (1,0) + (0,1) \overleftarrow{\times} (0,1)).$$

9. **Can quantum physics be made more complete? (Part II).** This Excursion concludes the discussion started by the Excursion of the same name in Week 1. We now include entanglement, to fix the hole that we left then. The experimental setup involves a source of polarized photons entangled in pairs and two polarization detectors whose angles can be altered at random *between the time that the photon pair is created and the time that the photons arrive at the detectors* so that no signal at lightspeed or less can tell one detector what the other one is doing. We'll use two angles for each detector and call them A, B, C and D (say $A = \pi/4$, $B = 0$, $C = \pi/8$ and $D = 3\pi/8$): A and B are the two angles the left detector can check and C and D are the two angles for the right detector. So the following four combinations are possible: AC, AD, BC and BD (say $-\pi/8$, $\pi/8$, $\pi/8$ and $3\pi/8$).
We'll use expected values of correlations between the polarizations, which can range from -1 to 1, instead of probabilities (as in Week 1), which range from 0 to 1.
a) I am following [Shi09] here but will be a little glib about the quantum physics. We write the entangled state (see the previous Excursion)

$$< A | = \frac{1}{\sqrt{2}}((1,0) \overleftarrow{\times} (1,0) + (0,1) \overleftarrow{\times} (0,1)).$$

and its transpose as

$$| A >= \frac{1}{\sqrt{2}} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

These bracket the operator which gives the angle between the two detectors, e.g., AB, but in tensor product, one applied to the second part of the state and the other to the first part. This operator is the 2D rotation matrix—but we remember to transpose it when applied to the left. With $c^2 + s^2 = 1$:

$$R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

$$\begin{aligned} < A \mid R^{\mathrm{T}(2)} R^{(1)} \mid A > &= \frac{1}{2} \left((1,0) \overleftarrow{\times} (c,s) + (0,1) \overleftarrow{\times} (-s,c) \right) \left(\begin{pmatrix} c \\ s \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -s \\ c \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ &= \frac{1}{2}(c^2 - s^2 - s^2 + c^2) \\ &= c_2 \end{aligned}$$

where c_2 is double the angle between the detectors.
For many repetitions the overall expected value for each case is given by

	$<AC>$	$<AD>$	$<BC>$	$<BD>$
θ	$-\frac{\pi}{8}$	$\frac{\pi}{8}$	$\frac{\pi}{8}$	$\frac{3\pi}{8}$
$\cos 2\theta$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$

b) For the entangled version of "key theory" (see Week 1) the Bell inequality is based on a relationship which holds for any four variables each ranging from -1 to 1. These fall on points on or in a 4D hypercube centred at the origin. We can represent its corners by a Karnaugh map (also used in Note 5 of Week 10) on which we record the correlations. It is clearer to make a separate Karnaugh map for each correlation and then combine them at the end in whatever way we will wind up finding useful.

AC

	D	-1		1	
B	$A \backslash C$	-1	1	-1	1
-1	-1	1	-1	1	-1
	1	-1	1	-1	1
1	-1	1	-1	1	-1
	1	-1	1	-1	1

AD

	D	-1		1	
B	$A \backslash C$	-1	1	-1	1
-1	-1	1	1	-1	-1
	1	-1	-1	1	1
1	-1	1	1	-1	-1
	1	-1	-1	1	1

BC

	D	-1		1	
B	$A \backslash C$	-1	1	-1	1
-1	-1	1	-1	1	-1
	1	1	-1	1	-1
1	-1	-1	1	-1	1
	1	-1	1	-1	1

BD

	D	-1		1	
B	$A \backslash C$	-1	1	-1	1
-1	-1	1	1	-1	-1
	1	1	1	-1	-1
1	-1	-1	-1	1	1
	1	-1	-1	1	1

What we are going to need is the sum of any three of these minus the fourth.

$AC + AD + BC - BD$

	D	-1		1	
B	$A \backslash C$	-1	1	$=1$	1
-1	-1	2	-2	2	-2
	1	-2	-2	2	2
1	-1	2	2	-2	-2
	1	-2	2	-2	2

We notice that the absolute value is always 2. So, if we allow intermediate values between -1 and 1, not just the corners of the hypercube, we have (writing $< AC >$ for the *expected* value of AC)

$$-2 \ \leq \ < AC > + < AD > + < BC > - < BD > \leq \ \ 2$$

c) But if we calculate the same expression using the expected values derived from quantum theory in part (a) we get $2\sqrt{2} = 2.83$ which violates the above inequality.
d) Again, not every set of angles for the detectors causes quantum theory to violate the inequality. The expression in part (a) is $3\cos 2\pi/8 - \cos 6\pi/8$. Plot $3\cos 2\theta - \cos 6\theta$ to see when the violation happens and where the maximum violation is.
e) Bell's original paper [Bel64] works with electrons. The quantum-theoretical calculation is similar to that in part (a). The entangled state is

$$< A \mid = \frac{1}{\sqrt{2}}((1,0) \overleftarrow{\times} (0,1) - (0,1) \overleftarrow{\times} (1,0)).$$

and its transpose is

$$\mid A > = \frac{1}{\sqrt{2}} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$$

The operator must rotate spin-1/2 particles in 3D, so it is the product of the vector of Pauli spin matrices with the orientation of the detector, again once fod each of the pair in the entangled state. Thus

$$\vec{\sigma} \cdot \vec{a} = \begin{pmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{pmatrix}$$

$$\begin{aligned} < A \mid (\vec{\sigma} \cdot \vec{b})^{\dagger(2)} (\vec{\sigma} \cdot \vec{a})^{(1)} \mid A > \ &= \ \frac{1}{2} \Big((1,0) \overleftarrow{\times} (b_x + ib_y, -b_z) - (0,1) \overleftarrow{\times} (b_z, b_x - ib_y) \Big) \\ &\quad \left(\begin{pmatrix} a_z \\ a_x + ia_y \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} a_x - ia_y \\ -a_z \end{pmatrix} \overleftarrow{\times} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \\ &= \ \frac{1}{2}(-2a_z b_z - 2a_x b_x - 2a_y b_y) \\ &= \ -\vec{a} \cdot \vec{b} \end{aligned}$$

10. a) For a normalized 3D vector (p, q, r), $p^2 + q^2 + r^2 = 1$. Think of a way of replacing p, q, r by cosines and sines. You'll need two angles: which angles are conventionally used?
b) Show that $(-q, p, 0)$ and $(p, q, -(p^2 + q^2)/r)$ could be the remaining two of a mutually orthogonal set of three vectors. How would they be normalized?
c) As a 2-by-1 vector (p, q, r) could be represented as $(p + iq, r)$. Is it normalized? Write two 2-by-1 vectors orthogonal to it. Hint: $i(p + iq) = ip - q$ is orthogonal to $p + iq$: what is the product of the two in a 2-by-2 vector product? (Remember to use the Hermitian transpose.)
d) Re-express the cos-sin answer to part (a) as a 2-by-1 as in part (c).

11. Show that f_3 does not generate extra ladder operators to give us the anticommutator algebra we seek in Note 19. E.g., $[D_1, D_2]_{(+)} \neq 0$, etc., in

$$\begin{aligned} D_1 \ &= \ \frac{1}{2}(f_2 + if_3) \\ D_2 \ &= \ \frac{1}{2}(f_3 + if_1) \\ D_3 \ &= \ \frac{1}{2}(f_1 + if_2) \end{aligned}$$

12. Work out all 16 elements of the $n = 4$ reflection given as f_0, f_1, f_2 and f_3 in Note 20. Work out the complete multiplication table and note which pairs commute and which anticommute.

13. Confirm that the two down and two up ladders generated in Note 20, from the $n = 4$ reflection algebra, satisfy the anticommutator relations given in Note 18.

14. Show that the three pairs of D and U operators arise from the 8-by-8 reflection algebra (based on Note 20):

$$\begin{aligned} f_1 &= f_z \overleftarrow{\times} I \overleftarrow{\times} I \\ f_2 &= f_x \overleftarrow{\times} f_y \overleftarrow{\times} I \\ f_3 &= f_x \overleftarrow{\times} f_x \overleftarrow{\times} f_x \\ f_4 &= f_x \overleftarrow{\times} f_x \overleftarrow{\times} f_y \\ f_5 &= f_x \overleftarrow{\times} f_x \overleftarrow{\times} f_z \\ f_6 &= f_x \overleftarrow{\times} f_z \overleftarrow{\times} I \end{aligned} \qquad \begin{aligned} D_1 &= f_1 + if_4 \\ D_2 &= f_2 + if_5 \\ D_3 &= f_3 + if_6 \\ U_1 &= f_1 - if_4 \\ U_2 &= f_2 - if_5 \\ U_3 &= f_3 - if_6 \end{aligned}$$

15. Check that the eigenvalues of J_x and J_y in Note 21 are also 1, 0 and -1, as are the eigenvalues of $pJ_x + qJ_y + rJ_z$ for any numbers p, q and r such that $p^2 + q^2 + r^2 = 1$. Hint. Use MATLAB's `eig()` function and symbolic variables `syms p,q,r`.

16. Show that if two matrices, M_1 and M_2, can be diagonalized by the same transformation X, then they commute. Make the slightly less easy converse argument.

17. Show that the 2-by-2 J_x, J_y and J_z in Note 21 have the same commutator relationships as the 3-by-3.

18. Unfortunately the exponential notation in Note 21 can be misleading.

$$e^{-i\theta_x J_x} e^{-i\theta_y J_y} e^{-i\theta_z J_z}$$

is not the same as

$$e^{-i(\theta_x J_x + \theta_y J_y + \theta_z J_z)}$$

Why? Hint. Think commutators.

19. Compare the 2-by-2 matrix resulting from using the Euler angles in Note 21 with Feynman's [FLS64, p.III-6-12]. Hint: Feynman is changing the *axes*; I've done the rotation with fixed axes.

20. Calculate the commutator/anticommutator table of Note 22. You can do it with just the rules of squaring and of swapping indices. Or, if this is tedious and error-prone, you can write a program: write out the three 2-by-2 reflections (call them `px`, `py`, `pz` for Pauli), use the MATLAB function `kron()` (e.g., given the 2-by-2 identity matrix, `f0 = kron(pz,I2)`) and matrix multiplication to generate the sixteen. Use a cell array to test the commutator and anticommutator in a double loop.

21. Feynman [Fey49, p.751] introduces what later becomes his "slash" notation (Note 22) without the slashes. It would be nice if eventually the slashes were not needed and we just used A, b, ∂, etc. But convention is against this and the slashes are unambiguous.

22. Work out the general spinors, from Note 22, in 3D and 4D (Minkowski). What are the vectors and spinors in 4D (Euclidean)?

23. Look up the Weyl, Majorana, Wedderburn, etc. representations of the 4D (16-element reflection (Clifford) algebra (Note 22)). Look up Weyl, Majorana, etc., spinors and relate them to these representations.

24. Would the matrix representation of a 5D, 32-element reflection algebra be unique, as is the (Pauli) representation of the 3D, 8-element reflection algebra?

25. Show that the distributive property of the tensor product $A \overleftarrow{\times} C + B \overleftarrow{\times} C = (A+B) \overleftarrow{\times} C$ (Note 23) follows from the distributive property of the numerical elements $ac + bc = (a+b)c$.

26. **Tensor product transposes.** Show that $(A \overleftarrow{\times} B)^T = A^T \overleftarrow{\times} B^T$.

27. Check the orthonormality of $\mid n_1 n_2 .. n_k .. >$ and $\mid n'_1 n'_2 .. n'_k .. >$ in Note 23.

28. Show that $e^{-ik\!\!\!/\cdot x\!\!\!/}$ is a solution to the Klein-Gordon equation of Note 24.

$$\begin{aligned} \partial_t^2 e^{-ik\!\!\!/\cdot x\!\!\!/} &= \partial_t(-i\omega_k e^{-ik\!\!\!/\cdot x\!\!\!/}) = -\omega_k^2 e^{-ik\!\!\!/\cdot x\!\!\!/} \\ \partial_x^2 e^{-ik\!\!\!/\cdot x\!\!\!/} &= \partial_x(ik_x e^{-ik\!\!\!/\cdot x\!\!\!/}) = -k_x^2 e^{-ik\!\!\!/\cdot x\!\!\!/} \end{aligned}$$

etc.

29. Find the first and second order (Note 26) corrections to the eigenvalue of the shear matrix (Note 19)

$$H_0 = \frac{1}{2}\begin{pmatrix} d+1/d & d-1/d \\ d-1/d & d+1/d \end{pmatrix}$$

with the eigenvectors

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and corresponding eigenvalues d and $1/d$ respectively (check these!),
a) if $H_I = I$, the identity matrix;
b) if $H_I = f_x = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$.

30. Show that the conjugate of the Dirac Lagrangian in Note 27

$$\begin{aligned} \mathcal{L}^\dagger &= (\overline{\psi}(i\ \partial\!\!\!/ - m)\psi)^\dagger \\ &= -\overline{\psi}(i\ \partial\!\!\!/ + m)\psi \end{aligned}$$

so that both components of the Klein-Gordon equation can be derived from the one Lagrangian.

31. **Hermitian transposes of reflections, etc.** For Note 27 we need Hermitian conjugates of 4-by-4 ladder operators and the d_μ.
a) Show that $d_j^\dagger = -d_j, j = 1,2,3$, using $d_j = ip_2 \overleftarrow{\times} p_j$ for the 2-by-2 Pauli matrices (in Note 23)

$$xp_x + yp_2 + zp_3 = \begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix}$$

b) Show that $d_0^\dagger = d_0$, where $d_0 = p_3 \overleftarrow{\times} I$
c) Show that $f_\mu^\dagger = f_\mu, \mu = 0,1,2,3$ where $f_0 = d_0$, $f_j = p_1 \overleftarrow{\times} p_j$. Hint. See Excursion *Tensor product transposes.*
d) Hence show that $D_{(+)}{}^\dagger = U_{(+)}$ for $D_{(+)} = f + if'$, $U_{(+)} = f - if'$ for any two reflections, f and f', from question (c).

32. For Note 28 show that the Cauchy-Riemann equations oblige $\bigtriangledown^2 u = 0 = \bigtriangledown^2 v$

33. **Slopes of Clifford functions.** Since f_1, f_2 are basic *reflections*, different from 1 and i which are basic rotations, maybe we would get results for functions of Clifford numbers (the reflection algebra) different from the Cauchy-Riemann equations of Note 28.
a) Show that this is not so. (You must decide on a convention for division, e.g., $f_1/f_2 = f_{12}$ (because $f_2^2 = 1$) rather than $f_2 \backslash f_1 = f_{21} = -f_{12}$.) Try $z = xf_1 + yf_2$ and $f = uf_1 + vf_2$
b) For Clifford numbers in 3D the "Clifford-Cauchy-Riemann" condition gives

$$\begin{pmatrix} \partial_x u & -\partial_x v & -\partial_x w \\ \partial_x v & \partial_x u & -\partial_y w \\ \partial_x w & \partial_y w & \partial_x u \end{pmatrix}$$

for the Jacobian. What is this?
c) What about the full 2D reflection (Clifford) algebra?

$$\begin{aligned} q &= w + xf_1 + yf_2 + zf_{12} \\ f(q) &= a + bf_1 + cf_2 + df_{12} \end{aligned}$$

34. Look up Augustin Louis Cauchy, 1789–1857. It is said he published a major paper every week in a peak phase: was this during his almost single-handed invention of the calculus of complex variables?

35. Make the argument in Note 29 that $\partial_t J_t + \partial_x J_x + \partial_y J_y + \partial_z J_z = 0$ for a current density in any direction, not just the z-direction. You may have to find a way of dividing throughout by δt rather than by Δz as the text does.

36. Using the Lagrangian of Note 29, show that $\Delta\mathcal{L}$ and $\not{\partial} \cdot \not{J}$ each go to zero.

37. Any part of the Prefatory Notes that needs working through.

References

[Bel64] John Bell. On the einstein podolsky rosen paradox. *Physics*, 1(3):195–200, 1964.

[Fey49] R. P. Feynman. The theory of positrons. *Phys. Rev.*, 76(6):749–59, Sept. 1949. URL=link.aps.org/doi/10.1103/PhysRev.76.749, last accessed 17/3/6.

[FLS64] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics, Volume III.* Addison-Wesley, 1964.

[Jor86] Thomas F Jordan. *Quantum Mechanics in Simple Matrix Form.* JohnWiley & Sons, Inc., New York, 1986.

[Shi09] Abner Shimony. Bell's theorem. URL https://plato.stanford.edu/entries/bell-theorem/ last accessed 18/7/6, 2009.