**Relativistic Quantum Mechanics** PhL 3.26.05\
**Bjorken and Drell**

Contents:


Chapter 1: The Dirac Equation (2)..... 2
1.1 Formulation of Relativistic Quantum Theory (2) ..... 2
1.2 Early Attempts (4)..... 2
1.3 The Dirac Equation (6) ..... 4
1.4 Nonrelativistic Correspondence (10) ..... 6
**History**..... 12
Chapter 2: Covariance of the Dirac Equation (16)..... 13
2.1 Covariant form of the Dirac Equation (16). ..... 13
2.2 Proof of Covariance (18)..... 14
2.3 Space Reflection (24)..... 24
2.4 Bilinear Covariants (25)..... 24
Chapter 3: Solutions of the Dirac Equation for a Free Particle (27)..... 24
3.1 Plane Wave Solutions (28)..... 24
3.2 Projection Operators for Energy and Spin ..... 33
3.3 Physical Interpretation of Free-particle Solutions and Packets (35)..... 40
Chapter 4: The Foldy-Wouthuysen Transformation (46) ..... 46
4.1 Introduction (46). ..... 46
4.1 Free-particle FW Transformation (46)..... 47
4.2 The FW Transformation with E&M Fields (48)..... 48
4.3 The Hydrogen Atom (52)..... 53
Chapter 5: Hole Theory (64)..... 54
5.1 The Problem of Negative-energy Solutions (64). ..... 54
5.2 Charge Conjugation (66)..... 55
5.3 Vacuum Polarization (70) ..... 57
5.4 Time Reversal and Other Symmetries (71)..... 57
Chapter 6: Propagator Theory (78) ..... 60
6.1 Introduction..... 60
6.2 The Nonrelativistic Propagator (71) ..... 60
6.3 Formal Definitions and Properties of Green's Functions (still non-relativistic situation)..... 65
6.4 The Propagator in Positron Theory (89). ..... 67
Chapter 7: Applications (100)..... **Error! Bookmark not defined.**
7.1 Coulomb Scattering of Electrons (100) ..... **Error! Bookmark not defined.**
Preliminary Comments: Compare to Rutherford result, and discuss BD units.**Error! Bookmark not defined.**
Phase Space:..... **Error! Bookmark not defined.**
Side Exercise with jμ. ..... **Error! Bookmark not defined.**
7.2 Trace Theorems and Final Evaluation of the Mott cross section (103)**Error! Bookmark not defined.**
7.3 Coulomb Scattering of Positrons (106)..... **Error! Bookmark not defined.**

"Bj" (James Bjorken) was born 1934, fud Stanford, and Prof there at SLAC. He is now 74 years old. Sidney Drell is now emeritus at SLAC and seems to have political interests, being in the Hoover and JASON and such things like Fred Zach did, did nuclear arms control work, is a violinist, Los Alamos board. Perhaps a person similar to Jim's buddy Bill Fraser, who by the way has a very low web profile,

> William Fraser, senior vice president for academic affairs at the University of California and vice chairman of CARA, talked about the funding of the telescope by Keck, and the founding of CARA. And keeping with tradition, Cronkite signed off with, "And that's the way it was."

CARA is California Association for Research in Astronomy.

Oddly, I never met Bjorken or Drell in my 7 years at UCB, though I probably attended talks one or the other gave.

**Preface**: They are taking the "propagator approach" pioneered by Feynman, heavy on the graphs, some think these *are* the theory. They would like experimentalists and students to understand all this stuff better. Reader must know QM at Schiff level.

A long list of things NOT covered in the book: the action/variational approach of Schwinger, axiomatic field theory (the mathematics), pure S-matrix theory (Chew), *bound states*, fancy dispersion relation stuff, massive vector mesons (too bad), derivative couplings. We are referred to a long list of books on these not-included subjects, on which we find my Mandl book and Chew's book. Since I was in the Chew world, I never even saw these books, but the names are dimly familiar.

## Chapter 1: The Dirac Equation (2)

**1.1 Formulation of Relativistic Quantum Theory (2)** . Authors write out 6 basic facts that we all know from regular quantum mechanics, and want to make sure these things are still true in the relativistic extension.

(1) there is a state vector $\psi$ in Hilbert space and prob is $|\psi|^2$ , function of $q_i$, $s_i$ and time t
(2) observables correspond to Hermitian operators, they write "hermitian".
(3) These operators have eigenstates with real eigenvalues $\omega_n$
(4) You can describe an arbitrary system state by an expansion on complete basis functions that probably come from the eigenfunctions of an eigenvalue equation.
(5) In such an expansion with coefficients $a_n$, prob of measuring $\omega_n$ is $|a_n|^2$.
(6) The SE tells how $\psi$ moves in time

**1.2 Early Attempts (4).** In non-rel, we have $H\psi = i\hbar\partial\psi/\partial t$ (and $H\psi^* = - i\hbar\partial\psi^*/\partial t$) and $H = p^2/2m$ for a free particle, and $\mathbf{p} = -i\hbar\nabla$, ie, we have the SE.

The first idea is to replace H with $\sqrt{p^2 + m^2}$ keeping $\mathbf{p} = -i\hbar\nabla$**,** hence 1.9. Authors claim that you then bring in all orders of the spatial derivative, and this makes you have a non-local theory which is a mess.

The second idea is to think $H^2 = \mathbf{p}^2 + m^2$ and then say $(i\hbar\partial_t)^2 \psi = (\mathbf{p}^2 + m^2)\psi$ so you are sort of saying $E^2\psi = H^2\psi$. This is really the **Klein-Gordon equation** $(\partial_\mu\partial^\mu + m^2)\psi = 0$ with $\hbar$=c=1 for the moment. What is wrong with this? To see,

We have to <u>first go back to regular quantum mechanics</u> with the regular SE. I am going to do this including a potential V which may be complex, anticipating a need to later derive 20.1 on page 130.
Suppose we define

$$\mathbf{j} = \hbar/(2im) [ \psi^*(\nabla\psi) - (\nabla\psi^*) \psi] = \hbar/(2im) [ \psi^*(\nabla\psi) - cc ] = \text{a "current"} \qquad (7.3)$$

Apply div on both sides to get

$$\nabla\bullet \mathbf{j} = \hbar/(2im) [(\nabla\psi^*) \bullet (\nabla\psi) + \psi^*\nabla^2\psi - cc] = \hbar/(2im) [ \psi^*\nabla^2\psi - cc]$$

since the two $(\nabla\psi^*) \bullet (\nabla\psi)$ terms cancel. Now since $H = p^2/2m + V = -\hbar^2\nabla^2/2m + V$, we replace

$$\nabla^2 = 2m(H-V)/(-\hbar^2) \quad \Rightarrow$$

$$\nabla\bullet \mathbf{j} = \hbar/(i) [ \psi^*(H-V)\psi - ([H-V]\psi )^*\psi]/(-\hbar^2) = (i/\hbar) [ \psi^*H\psi - (H\psi )^*\psi - \psi^* V \psi + V^* \psi\psi ]$$

$$= (i/\hbar) [ \psi^*H\psi - (H\psi )^*\psi - (V-V^*) \psi^* \psi]$$

$$= (i/\hbar) [ \psi^*H\psi - (H\psi )^*\psi - 2iIm(V)\psi^* \psi]$$

$$= (i/\hbar) [ \psi^*H\psi - (H\psi )^*\psi] + (2/\hbar) Im(V) \psi^*\psi$$

We still have that minus sign between the two terms, the second is still "cc" of the first. *But now* we use these facts:

$$H\psi = i\hbar\partial_t\psi \qquad (H\psi)^* = - i\hbar\partial_t\psi^*$$

and now that minus sign between the terms becomes a plus

$$\nabla\bullet \mathbf{j} = (2/\hbar) Im(V) \psi^*\psi = (i/\hbar) [ \psi^*( i\hbar\partial_t\psi ) + (i\hbar\partial_t\psi^* )\psi]/(-\hbar^2)$$

$$= (i/\hbar)(i\hbar) [ \psi^*( \partial_t\psi ) + (\partial_t\psi^* )\psi]$$

$$= - \partial_t(\psi^*\psi) = - \partial_t\rho \qquad \rho = \psi^*\psi = \text{positive!}$$

So we conclude that

$$\nabla\bullet \mathbf{j} + \partial_t\rho = (2/\hbar) Im(V) \rho \qquad \text{// agrees with 20.1 where } V_I \equiv - Im(V)$$

which, for a free particle or for real V, we can write as $\nabla\bullet\,\mathbf{j} + \partial_t\rho = 0$ or $\partial_\mu j^\mu = 0$ and we have a "conserved current". This is of course a probability current, not an electric current, and ρ is the probability density. The key idea is that ψ*ψ is positive definite so you can interpret it as a probability. And

$j_\mu = \hbar/(2im)\,[\,\psi^*(\partial_\mu\psi) - (\partial_\mu\psi^*)\,\psi]$ = conserved 4-current

Now if we try using this same current in the $E^2\psi = H^2\psi$ theory we run into a problem. We get

$\rho = j_0 = \hbar/(2im)\,[\,\psi^*(\partial_t\psi) - (\partial_t\psi^*)\,\psi]$

which is the same as before in terms of the wavefunctions. But NOW we don't have a way to show that ρ is positive! We can no longer say $\partial_t\psi = +(1/i\hbar)\,H\psi$ and $\partial_t\psi^* = -(1/i\hbar)\,H\psi^*$, so we cannot get rid of that minus sign!

When history saw this problem, it put aside the Klein-Gordon equation and moved to the Dirac equation. The KG paper was in 1927, and Dirac was 1928. It would be fascinating to read these papers, but not now!

**1.3 The Dirac Equation (6)** I never appreciated before what Dirac did, it is such a simple idea. We want to get back to the linear time dependence so we have $H\psi = i\hbar\partial\psi/\partial t$ and we can maybe rescue the above probability problem. So try adding α and β *matrix* coefficients as shown in 1.13. Scalar coefficients fail because you are not rotationally covariant, not clear what matrices do on that, but we shall see later. So assume some kind of column vector solution and some NxN matrices for the $\alpha_i$ and β coefficients. In order to maintain the requirement that $E^2\psi = H^2\psi = (\mathbf{p}^2c^2 + m^2c^4)\psi = ([-i\hbar\nabla]^2c^2 + m^2c^4)\psi$, we require that the matrices have the properties shown in 1.16. This makes the second last term in equation A be zero, and makes the first give the $\nabla^2$ term and the last give the $m^2c^4$term. Very excellent, Mr. Dirac.

We require also that α and β be Hermitian matrices to keep H Hermitian, and then easy to show that the $\alpha_i$ are traceless. Since $\alpha^2 = \beta^2 = 1$, only eigenvalues can be ±1 (remember they are Hermitian so eigenvalues must be real). In order to make the diagonalized $\alpha_i$ have trace 0, it must have an equal number of +1 and -1 eigenvalues, so N must be even!

Why not try N=2? If we try $\alpha_k = \sigma_k$, the first item in 1.16 is happy. But what can you use for β? Need a matrix such that $\beta\alpha_i = -\alpha_i\beta$. The entire space of matrices in N=2 is spanned by **σ** and 1 so

$\beta = \mathbf{a}\bullet\boldsymbol{\sigma} + b\,1$

where all coefficients can be complex, so there are 8 real numbers apropo for 2x2. So we need:

$[\,\mathbf{a}\bullet\boldsymbol{\sigma} + b\,]\,\sigma_k + \alpha_k[\,\mathbf{a}\bullet\boldsymbol{\sigma} + b\,] = 0 \quad \Rightarrow \quad a\bullet[\,\boldsymbol{\sigma}\sigma_k + \sigma_k\boldsymbol{\sigma}] + 2b\sigma_k = 0$

$\Rightarrow a_k 1 + b\sigma_k = 0$

But $tr(\sigma_k) = 0$ so this says that $2a_k = 0$ so we end up with $a_k = 0$ and then b = 0 as the only solution! You could mount a completely formal proof, but this is enough for me, there is no N=2 solution.

There *is* an N=4 solution and it is given by 1.17. I have verified on scratch that 1.16 are all true and the matrices are Hermitian, as required.

Page 9 then constructs a conserved current with a ρ that is positive! Recall that our non rel current that worked was this:

$$j_\mu = \hbar/(2im)\,[\,\psi^*(\partial_\mu\psi) - (\partial_\mu\psi^*)\,\psi] \qquad \partial_\mu j^\mu = 0 \qquad \rho = \psi^*\psi$$

The current that works in the Dirac theory does not have gradients, it is just this:

$$\mathbf{j} = c\,\psi^\dagger \boldsymbol{\alpha}\,\psi$$
$$\rho = \psi^\dagger\,\psi$$

Let's just do this out the way we did in the non-rel case:

$$\nabla\bullet\,\mathbf{j} = c\,[(\nabla\psi^\dagger)\bullet\boldsymbol{\alpha}\,\psi + \psi^\dagger\boldsymbol{\alpha}\bullet(\nabla\psi)] = ?$$

Our Ham equation says

$$i\hbar\partial_t\psi = (\hbar c/i)\,\boldsymbol{\alpha}\bullet\nabla\psi + \beta mc^2\psi = H\psi$$
$$-i\hbar\partial_t\psi = -(\hbar c/i)\,\boldsymbol{\alpha}\bullet\nabla\psi^\dagger + \beta mc^2\psi^\dagger = H\psi^\dagger$$

which says that

$$(\hbar c/i)\,\boldsymbol{\alpha}\bullet\nabla\psi = i\hbar\partial_t\psi - \beta mc^2\psi$$
$$(\hbar c/i)\,\boldsymbol{\alpha}\bullet\nabla\psi^\dagger = i\hbar\partial_t\psi + \beta mc^2\psi$$

$$c\boldsymbol{\alpha}\bullet\nabla\psi = -\partial_t\psi - i(\beta mc^2)/h\,\psi$$
$$c\boldsymbol{\alpha}\bullet\nabla\psi^\dagger = -\partial_t\psi^\dagger + i(\beta mc^2)/h\,\psi^\dagger$$

Therefore

$$\nabla\bullet\,\mathbf{j} = c\,[(\nabla\psi^\dagger)\bullet\boldsymbol{\alpha}\,\psi + \psi^\dagger\boldsymbol{\alpha}\bullet(\nabla\psi)] = -(\partial_t\psi^\dagger)\psi - \psi^\dagger(\partial_t\psi) = -\partial_t[\psi^\dagger\psi] \quad \text{QED}$$

So somehow $c\boldsymbol{\alpha}$ is a "current operator" in this Dirac world. Of course $\psi^\dagger\psi = |\psi|^2$ since we are in a metric space where $(\psi,\psi) = \psi^\dagger\psi$ in the usual matrix fashion, and we can interpret ρ again as a probability.

So what has Dirac just done here?

(1) $i\hbar\partial_t\psi = (\hbar c/i)\,\boldsymbol{\alpha}\bullet\nabla\psi + \beta mc^2\psi = H\psi$ is our candidate new Schrodinger equation
which just means that $H = (\hbar c/i)\,\boldsymbol{\alpha}\bullet\nabla + \beta mc^2$ is our candidate Hamiltonian. So we use the same old SE, but ψ is now a 4-component column vector of functions and H is a 4x4 matrix.

(2) This form gives the result $H^2\psi = E^2\psi = (i\hbar\partial_t)^2\psi = (\mathbf{p}^2c^2 + m^2c^4)\psi$. In other words, we get that

$H^2\psi = ([-i\hbar\nabla]^2c^2 + m^2c^4)\psi$ . In other words, $H^2$ is what it must be to maintain the $\mathbf{p} = -i\hbar\nabla$ relationship and have the right notion that $H^2 = E^2$, relativistically correct.

(3) We get a conserved current with $\rho = \psi^\dagger \psi \geq 0$ as required.

The one missing thing we have not shown is that somehow the SE is covariant with this H! I guess this is complex enough to show that B&D have an entire chapter devoted to it, Chapter 2.

Another item is this: we have not said what kind of "particle" this Dirac Hamiltonian might work for?

I wonder how excited Dirac was when he figured this out. It is something I think I could have done in the same circumstance, nothing too messy really.

I think these $\psi$ 4-vectors belong to the $\frac{1}{2} \oplus 1 + 1 \oplus \frac{1}{2}$ representation of the Poincare Group, but that is a long distance away right now. I wonder what was known about such group representations in the Dirac timeframe of 1928 when he did this work?

**1.4 Nonrelativistic Correspondence (10)**. This will be the big test of the Dirac idea: do we get anything resembling the real world in the NR limit of the Dirac theory??

Before I can do notes for this chapter, we have to pre-derive some equations which are extremely important to the discussion, and which are very non-obvious. So let's do that now:

---

Aside on equations 1.27: There is a lot packed in here, a typical B&D-ism that requires a 2 hour digression on the part of the reader.

In the Schrodinger Picture, operators are fixed and the states move under the force of H.

$|\alpha(t)> = e^{-iHt/\hbar} |\alpha(0)$ H // Schiff page 169 24.3

In the Heisenberg Picture, operators move according to the commutator relation and states are fixed.

$dQ/dt = (i/\hbar)[H,Q] + \partial Q/\partial t$ $|\alpha>$ // Schiff page 169 24.4 or 24.10 p 171

In the Interaction Picture with $H = H_0 + H'$, states move with H', operators with $H_0$.

$dQ/dt = (i/\hbar)[H_o,Q] + \partial Q/\partial t$ $|\alpha(t)> = e^{-iH't/\hbar} |\alpha(0)$ // Schiff 24.14 and 24.13 approx

Now, with regard to 1.27, what "picture" is implied? It cannot be the Schrodinger picture because in that picture, you don't have an H commutator relation at all, so it must be one of the other two pictures. In the text, we are told to think of equation 1.26 as saying $H = H_0 + H'$. Defining the H's in this way, we conclude that 1.27 must refer to the *Heisenberg Picture* since the "full H" appears in the commutators.

1.27 #1: The commutator $[H,\mathbf{r}]$ is nonzero only due to the p item in H, so we have

$[H,r_i] = c\boldsymbol{\alpha}\bullet[\mathbf{p},r_i] = c\alpha_j[p_j, r_i] = -i\hbar c\alpha_i = (\hbar/i)\, c\alpha_i$ since $[p,x] = -i\hbar$

Therefore, $dr_i/dt = (i/\hbar)[H,r_i] = c\alpha_i$ or $d\mathbf{r}/dt = c\boldsymbol{\alpha} \equiv \mathbf{v_{op}}$

So we have now proven this first result and defined $\mathbf{v_{op}}$ as shown. We are seeing that $c\boldsymbol{\alpha}$ is the correct velocity operator in this theory. Notice that operator **r** has no "explicit" time dependence so we did not have any term of the form $\partial\mathbf{r}/\partial t$.

1.27 #2: We start of here saying:

$$d\boldsymbol{\pi}/dt = (i/\hbar)[H,\boldsymbol{\pi}] + \partial\boldsymbol{\pi}/\partial t \qquad \boldsymbol{\pi} = \mathbf{p} - (e/c)\mathbf{A}(\mathbf{r},t)$$

In Hamiltonian theory, **r** and **p** are the arguments of H, and neither has "explicit time dependence", so as before we have $\partial\mathbf{p}/\partial t = 0$. This means $\partial\boldsymbol{\pi}/\partial t = -(e/c)\,\partial_t\mathbf{A}$ . Thus we have shown that

$$d\boldsymbol{\pi}/dt = (i/\hbar)[H,\boldsymbol{\pi}] - (e/c)\,\partial_t\mathbf{A}$$

which verifies the second equation

1.27 #3: Let's compute $[H, \pi_i] = [H, p_i] - (e/c)\,[H, A_i(\mathbf{r})]$. Here we show the r dependence of A as a reminder that this is not going to commute with p-things. So the first term is this

$$[H, p_i] = [\, c\boldsymbol{\alpha} \bullet (\mathbf{p} - [e/c]\mathbf{A}(\mathbf{r})) + \beta mc^2 + e\varphi(\mathbf{r}),\, p_i]$$

$$= -e\,\boldsymbol{\alpha} \bullet [\mathbf{A}(\mathbf{r}),\, p_i] + e\,[\varphi(\mathbf{r}),\, p_i] \qquad \text{// now use rule } [f(\mathbf{r}),p_i] = +\,i\hbar\,\partial_i f(\mathbf{r})$$

$$= -i\hbar e\,\{\alpha_j\,\partial_i A_j(\mathbf{r}) - \partial_i\varphi\,\}$$

In vector form, this says

$$[H, \mathbf{p}] = -i\hbar e\,\{\alpha_j\,\nabla A_j(\mathbf{r}) - \nabla\varphi\,\} = -i\hbar e\,\nabla\,\{(\boldsymbol{\alpha}\bullet\mathbf{A}(\mathbf{r})) - \varphi\,\}$$

where the gradient is the vector sense in each term on the RHS.

Meanwhile, we now have to compute

$$[H, A_i(\mathbf{r})] = [\, c\boldsymbol{\alpha} \bullet (\mathbf{p} - [e/c]\mathbf{A}(\mathbf{r})) + \beta mc^2 + e\varphi(\mathbf{r}),\, A_i(\mathbf{r})]$$
$$= c\boldsymbol{\alpha} \bullet [\mathbf{p},\, A_i(\mathbf{r})] = -i\hbar\; c\boldsymbol{\alpha} \bullet \nabla A_i(\mathbf{r}) \qquad \text{// used } [\mathbf{p},\, f(\mathbf{r})\,] = -i\hbar\nabla f(r)$$
$$= -i\hbar c\,\alpha_j\partial_j A_i$$

Therefore we have

$$[H, \pi_i] = [H, p_i] - (e/c)\,[H, A_i] = -i\hbar e\,\{\alpha_j\,\partial_i A_j(\mathbf{r}) - \partial_i\varphi\,\} + i\hbar e\,\alpha_j\partial_j A_i$$

$$= -i\hbar e\,\{\, -\partial_i\varphi + \alpha_j\,\partial_i A_j - \alpha_j\partial_j A_i\,\} = -i\hbar e\,\{\, -\partial_i\varphi + \alpha_j\,(\partial_i A_j - \partial_j A_i)\,\}$$

But write $(\partial_i A_j - \partial_j A_i) = \varepsilon_{ijk} [\nabla \times A]_k = \varepsilon_{ijk} B_k$ . For each choice of $i \neq j$, there is only one surviving term in the implied k sum, otherwise there are no terms. Then we have

$$\alpha_j \ (\partial_i A_j - \partial_j A_i) = \alpha_j \ \varepsilon_{ijk} B_k = \varepsilon_{ijk} \ \alpha_j B_k = [\ \boldsymbol{\alpha} \times \mathbf{B}\ ]_i$$

and then

$$[H, \pi_i] = -i\hbar e \ \{ \ - \partial_i \varphi + [\ \boldsymbol{\alpha} \times \mathbf{B}\ ]_i \ \} = (\hbar/i) e \ \{ \ - \partial_i \varphi + [\ \boldsymbol{\alpha} \times \mathbf{B}\ ]_i \ \}$$

Now we use 1.27 #2 which said

$$d\boldsymbol{\pi}_i/dt = (i/\hbar)[H,\boldsymbol{\pi}_i] - (e/c) \ \partial_t A_i = e \ \{ \ - \partial_i \varphi + [\ \boldsymbol{\alpha} \times \mathbf{B}\ ]_i \ \} - (e/c) \ \partial_t A_i$$

$$= e \ \{ \ - \partial_i \varphi - \ (1/c) \ \partial_t A_i \ + [\ \boldsymbol{\alpha} \times \mathbf{B}\ ]_i \ \} = e \ \{ E_i + \ [\ \boldsymbol{\alpha} \times \mathbf{B}\ ]_i \ \}$$

or in vector form

$$d\boldsymbol{\pi}/dt = e \ \{ \ \mathbf{E} + \ \boldsymbol{\alpha} \times \mathbf{B} \ \} \qquad \text{and recall } c\boldsymbol{\alpha} \equiv \mathbf{v_{op}} \qquad \text{hence \#3 is proven.}$$

Now let's look again at these three results:

#1: $d\mathbf{r}/dt = c\boldsymbol{\alpha} \equiv \mathbf{v_{op}}$

#2: $d\boldsymbol{\pi}/dt = (i/\hbar)[H,\boldsymbol{\pi}] - (e/c) \ \partial_t \mathbf{A}$

#3: $d\boldsymbol{\pi}/dt = e \ \{ \ \mathbf{E} + \ \boldsymbol{\alpha} \times \mathbf{B} \ \} = e \ \{ \ \mathbf{E} + \ (1/c) \ \mathbf{v_{op}} \times \mathbf{B} \ \}$

The first says what "velocity" is in this Dirac theory. The second is intermediate. The third says what "force" is in this theory, and it is exactly the Lorentz Force where we must use the appropriate velocity.

Notice that all of these are matrix equations, though this is perhaps not obvious. For example, #1 says

$$d\mathbf{r}/dt \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = c \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}$$

where 1 is the 2x2 identity and $\boldsymbol{\sigma}$ are the Pauli 2x2's. The #3 equation says

$$d\boldsymbol{\pi}/dt \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = e\mathbf{E} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + e \begin{pmatrix} 0 & \boldsymbol{\sigma} \times \mathbf{B} \\ \boldsymbol{\sigma} \times \mathbf{B} & 0 \end{pmatrix}$$

---

Aside on Equation 1.33: (another non-obvious fact)

The problem here is to show this :

$i\,\boldsymbol{\pi} \times \boldsymbol{\pi}\,\varphi = (-e\hbar/c)\,\mathbf{B}\,\varphi$

or

$i\,(\mathbf{p} - (e/c)\mathbf{A}) \times (\mathbf{p} - (e/c)\mathbf{A})\,\varphi = (-e\hbar/c)\,\mathbf{B}\,\varphi$

or

$-i(e/c)\,\{\,\mathbf{A} \times \mathbf{p} + \mathbf{p} \times \mathbf{A}\,\}\varphi = (-e\hbar/c)\,\mathbf{B}\,\varphi$

or

$i\{\,\mathbf{A} \times \mathbf{p} + \mathbf{p} \times \mathbf{A}\,\}\varphi = \hbar\,\mathbf{B}\,\varphi$

One must remember that $\mathbf{p} = -i\hbar\nabla$, and in the second term in {...} the $\nabla$ is *acting on everything to its right*, it is not just acting on **A**. So throw this in:

$i\,(-i\hbar)\,\{\,\mathbf{A} \times \nabla + \nabla \times \mathbf{A}\,\}\varphi =?= \hbar\,\mathbf{B}\,\varphi$

or

$\{\,\mathbf{A} \times \nabla + \nabla \times \mathbf{A}\,\}\varphi =?= \mathbf{B}\,\varphi$

So if we can show this last line, we have it. Still the second term requires disambiguation. We can write

$\{\,\nabla \times \mathbf{A}\,\}_k\,\varphi = \varepsilon_{kij}\partial_i A_j\varphi$ means $\varepsilon_{kij}\partial_i(A_j\varphi) = \{\,\nabla \times (A\varphi)\,\}_k$

or

$\nabla \times \mathbf{A}\,\varphi$ means $\nabla \times (A\varphi)$

So here then is what we have to show:

$\mathbf{A} \times (\nabla\varphi) + \nabla \times (\mathbf{A}\varphi) =?= \mathbf{B}\varphi$ of course $\mathbf{B} = \nabla \times \mathbf{A}$

But I have a vector identity which says this

$\nabla \times (\mathbf{A}\varphi) = -A \times (\nabla\varphi) + (\nabla \times \mathbf{A})\,\boldsymbol{\varphi}$

so putting this in the previous equation, the $\mathbf{A} \times (\nabla\varphi)$ terms cancel and we get

$(\nabla \times \mathbf{A})\,\varphi =?= \mathbf{B}\varphi$

which is correct, QED.

---

Aside on Equation 1.35: (another non-obvious fact)

This claims to just be a rewrite of 1.34.

(1) Show that $\mathbf{A} = \tfrac{1}{2}\,\mathbf{B} \times \mathbf{r}$ is acceptable for a uniform **B** field, that is, show $\nabla \times \mathbf{A} = \mathbf{B}$.

$\nabla \times \mathbf{A} = \tfrac{1}{2}\,\nabla \times (\mathbf{B} \times \mathbf{r}) = \tfrac{1}{2}\,\{\,\mathbf{B}\,(\nabla \bullet \mathbf{r}) - (\mathbf{B} \bullet \nabla)\mathbf{r}\,\}$

which follows from the usual vector identity where two terms vanish because **B** = constant in **r**. Now

$(\nabla \bullet \mathbf{r}) = \partial_i r_i = 1 + 1 + 1 = 3$

$(\mathbf{B} \bullet \nabla) r_k = B_i \partial_i r_k = B_i \delta_{ik} = B_k$ so $(\mathbf{B} \bullet \nabla)\, \mathbf{r} = \mathbf{B}$

Thus,

$\frac{1}{2} \{ \mathbf{B} (\nabla \bullet \mathbf{r}) - (\mathbf{B} \bullet \nabla)\mathbf{r} \} = \frac{1}{2} \{ 3\mathbf{B} - \mathbf{B} \} = \mathbf{B}$ QED

(2) Compute $A^2$

$\mathbf{A} \bullet \mathbf{A} = \frac{1}{2}\, \mathbf{B} \times \mathbf{r} \bullet \frac{1}{2}\, \mathbf{B} \times \mathbf{r} = \frac{1}{4} \{ (\mathbf{B} \bullet \mathbf{B})(\mathbf{r} \bullet \mathbf{r}) - (\mathbf{B} \bullet \mathbf{r})^2 \} = \frac{1}{4} \{ B^2 r^2 - (\mathbf{B} \bullet \mathbf{r})^2 \}$

But we are going to ignore this term because it has $1/c^2$ compared to the linear term.

(3) The cross term in the $\pi^2$ part of H is this, where we have the same "to the right" issue we had in our previous derivations above, so we write this as

$-(e/c) \{ \mathbf{p} \bullet \mathbf{A} + \mathbf{A} \bullet \mathbf{p} \} \varphi = +(i\hbar e/c) \{ \nabla \bullet \mathbf{A} + \mathbf{A} \bullet \nabla \} \varphi$

$= (i\hbar e/c) \{ \nabla \bullet (\mathbf{A}\varphi) + \mathbf{A} \bullet (\nabla \varphi) \}$

But vector identity says

$\nabla \bullet (\mathbf{A}\varphi) = \mathbf{A} \bullet (\nabla \varphi) + (\nabla \bullet \mathbf{A})\, \varphi$ // where second $\nabla$ acts only on A

But

$\nabla \bullet \mathbf{A} = \frac{1}{2} \nabla \bullet (\mathbf{B} \times \mathbf{r}) = \frac{1}{2} \{ \mathbf{r} \bullet \nabla \times \mathbf{B} - \mathbf{B} \bullet \nabla \times \mathbf{r} \}$

but of course $\nabla \times \mathbf{B} = \mathbf{0}$ and $(\nabla \times \mathbf{r})_k = \varepsilon_{kij} \partial_i r_j = \varepsilon_{kij} \delta_{ij} = 0$ as well. so $\nabla \bullet \mathbf{A} = 0$. But I knew that because $\nabla \bullet (\nabla \times \mathbf{A}) = 0$ for any A, reminding us of Maxell $\nabla \bullet \mathbf{B} = 0$. OK, fine. So we then have

$\nabla \bullet (\mathbf{A}\varphi) = \mathbf{A} \bullet (\nabla \varphi)$

and the cross term in the $\pi^2$ part of H is this,

$= (i\hbar e/c) \{ \nabla \bullet (\mathbf{A}\varphi) + \mathbf{A} \bullet (\nabla \varphi) \} = (2i\hbar e/c)\ \mathbf{A} \bullet (\nabla \varphi)$

$= (2i\hbar e/c)\ \frac{1}{2}\, (\mathbf{B} \times \mathbf{r}) \bullet (\nabla \varphi) = -(e/c)\ (\mathbf{B} \times \mathbf{r}) \bullet (\mathbf{p}\varphi)$

$= -(e/c)\, \mathbf{B} \bullet \mathbf{r} \times (\mathbf{p}\varphi) = -(e/c)\, \mathbf{B} \bullet \mathbf{L}\, \varphi$

We have to divide by 2m from the first term in 1.34 so we get

contribution from $\pi^2$ term in H = - (e/2mc) $\mathbf{B} \bullet \mathbf{L}\ \varphi$

We have no electric potential in our uniform B field, so drop last term. And then

$\mathbf{S} \equiv (\hbar/2)\ \boldsymbol{\sigma}$

so the second term is then

$-(e\hbar/2mc)(2/\hbar)\mathbf{S} = -(e/mc)\mathbf{S} = -(e/2mc)2\mathbf{S}$

So the contribution of the above terms in H is then this:

$-(e/2mc)\ \mathbf{B} \bullet \mathbf{L}\ \varphi \quad -(e/2mc)2\mathbf{S} \bullet \mathbf{B}\ \varphi \quad = -(e/2mc)\ (\ \mathbf{L} + 2\mathbf{S}\ ) \bullet \mathbf{B}\ \varphi$

and this concludes our derivation of 1.35.

---

Finally, we may now begin our notes on this section! ("Nonrelativistic Correspondence")

The first item considered is "the electron at rest". So suddenly our authors are saying that this Dirac theory applies to electrons, maybe. The argument is made that, for such an electron, we can ignore the spatial gradient terms in H shown in 1.13. Yes, we know that $\lambda = h/p$ and as $p \rightarrow 0$ $\lambda \rightarrow \infty$ for the "wave" that describes an electron. I guess I buy this argument.

Then I agree that the four spinors shown in 1.24 are the correct independent rest solutions, and be sure to notice that the exponents have opposite sign in the last two solutions, these are the "negative energy solutions" as Mandl also likes to call them, due to the + sign in the exponent.

Now what comes next? I think the claim is that that prescription 1.25 $p_\mu \rightarrow \pi_\mu$ for bringing in the EM potentials is relativistically exact. At least we know it is a covariant thing to do, so let's do it. This makes our H as in 1.26 where we have now separated the space and time parts of the potential we just added.

The group of equations I call 1.27 are now interesting, but not quite in the main line of the discussion. First, these are Heisenberg Picture operator motion equations and I have derived them all above. There are two main points to learn here:

(1) the correct relativistic velocity operator seems to be the matrix $c\boldsymbol{\alpha} \equiv \mathbf{v}_{op}$

(2) the "Lorentz force" is really $d\boldsymbol{\pi}/dt$, not $d\mathbf{p}/dt$, and is given by the "usual formula" where we use the correct relativistic velocity $c\boldsymbol{\alpha}$ . I presume when we take the NR limit of this formula, we will recover the traditional formula somehow for Lorentz force, but BD don't do that here.

(3) Notice that the conserved space current is this : $\mathbf{j} = \psi^\dagger \mathbf{v}_{op} \psi$ from 1.22, $j_0 = \psi^\dagger \psi$. The velocity is the velocity of probability "flow".

Now we start afresh on page 12. We break the 4-vector into a pair of 2-vectors $\tilde{\varphi}$ and $\tilde{\chi}$, and we write the Dirac equation exactly as in 1.28. The flipping of the second spinor arises from the off-diagonal appearance of $\boldsymbol{\sigma}$ in the $\boldsymbol{\alpha}$ matrix as shown in 1.17 page 8.

At this point, we pretend that all solutions are positive energy solutions, and we remove that fast time dependence from the barred spinors (same for each spinor) thus defining unbarred ones, as in 1.29. Then 1.30 are the exact equations for the unbarred ones.

If you look at the second equation and ignore the two terms that don't have power of c in them, you get 1.31 which says that the $\chi$ spinor is smaller than the $\varphi$ spinor by v/c, so the $\varphi$ is the one we want to look at for our NR limit. Using 1.31 to eliminate $\chi$ from the first equation in 1.30 [ where $\chi$ appears in one place] , we get 1.32. Notice that, although $\chi$ is small, it gets boosted in the first equation by c so it should not be ignored, and it is the cause of the first term on the RHS of 1.32.

At this point, we do lots of algebra as outlined above, and we get 1.34 to drop out. Suddenly a "spin term" has appeared out of thin air, it is the $\boldsymbol{\sigma}\bullet\mathbf{B}$ term and this is saying that there is some kind of intrinsic magnetic moment proportional to $\sigma$ on this electron in the NR limit.

We then assume a uniform **B** field and get 1.35 which clinches things. This says that, in the NR limit of the Dirac theory, the "interaction Hamiltonian" involves the famous result

$$H' = -(e/2mc)\,(\mathbf{L} + 2\mathbf{S})\bullet\mathbf{B} \qquad \text{where } \mathbf{S} = (\hbar/2)\boldsymbol{\sigma} \text{ and } \mathbf{L} = \mathbf{r} \times \mathbf{p}$$

If we take the NR Hamiltonian $p^2/2m$ and make the spatial $\boldsymbol{\pi}$ replacement and keep the cross terms, we get the **L** term shown above, but we don't get the **S** term unless we add it "manually" by assuming that the electron has a certain magnetic moment. Here we have *derived* the presence of the second term from the Dirac equation. By finding a covariant $H\psi = E\psi$ formulation (though we have not yet shown covariance), and by looking at its solutions in the NR limit, the "spin" has appeared out of nowhere.

What order in history did things happen? I think the Dirac work above was 1928. The Stern-Gerlach experiment was 1922. Below is some spin history from wiki. Pauli made up the idea of a 2-valued but unknown internal quantum number "up/down" and that each spatial state could have one electron of each internal value -- This was the Pauli Exclusion Principle, they are saying 1924 for this. Pauli rejected Kronig's suggestion that this internal number related to angular momentum of the electron, but the Gerlach thing sure must have suggested that notion. Then in 1925 our friends U and G and E did some theory on this spin idea. Then Pauli changed his mind, and formalized the theory in 1927 and named the "Pauli matrices" and "2-spinors". Then in 1928 Dirac did his relativistic H and saw this spin stuff just pop out! That was a major physics victory of the 1920's era. A lot of stuff happened in a small time. Remember that the SE itself was only found 1926, and Heisenberg's matrix stuff was 1925-6 as well.

**History**

Spin was first discovered in the context of the emission spectrum of alkali metals. In 1924 Wolfgang Pauli introduced what he called a "two-valued quantum degree of freedom" associated with the electron in the outermost shell. This allowed him to formulate the Pauli exclusion principle, stating that no two electrons can share the same quantum state at the same time.

The physical interpretation of Pauli's "degree of freedom" was initially unknown. Ralph Kronig, one of Landé's assistants, suggested in early 1925 that it was produced by the self-rotation of the electron. When Pauli heard about the idea, he criticized it severely, noting that the electron's hypothetical surface would have to be moving faster than the speed of light in order for it to rotate quickly enough to produce the necessary angular momentum. This would violate the theory of relativity. Largely due to Pauli's criticism, Kronig decided not to publish his idea.

In the fall of 1925, the same thought came to two Dutch physicists, George Uhlenbeck and Samuel Goudsmit. Under the advice of Paul Ehrenfest, they published their results. It met a favorable response, especially after Llewellyn Thomas managed to resolve a factor of two discrepancy between experimental results and Uhlenbeck and Goudsmit's calculations (and Kronig's unpublished ones). This discrepancy was due to the orientation of the electron's tangent frame, in addition to its position. Mathematically speaking, a fiber bundle description is needed. The tangent bundle effect is additive and relativistic; that is, it vanishes if *c* goes to infinity. It is one half of the value obtained without regard for the tangent space orientation, but with opposite sign. Thus the combined effect differs from the latter by a factor two (Thomas precession).

Despite his initial objections, Pauli formalized the theory of spin in 1927, using the modern theory of quantum mechanics discovered by Schrödinger and Heisenberg. He pioneered the use of Pauli matrices as a representation of the spin operators, and introduced a two-component spinor wave-function.

Pauli's theory of spin was non-relativistic. However, in 1928, Paul Dirac published the Dirac equation, which described the relativistic electron. In the Dirac equation, a four-component spinor (known as a "Dirac spinor") was used for the electron wave-function. In 1940, Pauli proved the *spin-statistics theorem*, which states that fermions have half-integer spin and bosons integer spin.

# Chapter 2: Covariance of the Dirac Equation (16)

### 2.1 Covariant form of the Dirac Equation (16).

---

**Notation**: BD use $g_{\mu\nu} = g^{\mu\nu} = \text{diag}(1,-1,-1,-1)$. This makes $p_\mu p^\mu = + (mc)^2$, for example. I remember from somewhere this fact:

$\partial/\partial x^\mu = \partial_\mu$ that is, the thing on the left is really a lower index thing (covariant vector)

Now consider this:

$p^\mu = + i\hbar\, \partial/\partial x_\mu = +i\hbar\, \partial^\mu$

Then

$p^i = +i\hbar\, \partial^i = - i\hbar\, \partial_i = -i\hbar\, \nabla_i = -i\hbar\, \partial/\partial x^i$

$\mathbf{p} = -i\hbar\, \nabla$ // which agrees with the traditional formula

$p^0 = +i\hbar\, \partial^t = +i\hbar\, \partial_t$ // sign same for time component

Now consider

$\gamma^\mu \partial_\mu = \gamma^0 \partial_0 + \gamma^i \partial_i = \gamma^0 \partial_0 + \gamma^i \nabla_i = \gamma^0 \partial_0 + \boldsymbol{\gamma} \bullet \nabla$

Aside on the LT in contra/covariant notation. It is hard to find a source on this, but Messiah has it pretty well on page 877. The names are $x^\mu$ = contravariant vector, $x_\mu$ = covariant vector, and you use $g_{\mu\nu}$ to lower an index, and $g^{\mu\nu}$ to raise an index. So, suppose we define $g_{\mu\nu} = g^{\mu\nu} = \text{diag}(1,-1,-1,-,1)$. Then:

$g^{\alpha}{}_{\beta} = g^{\alpha\gamma} g_{\gamma\beta}$ because we are raising the first index $\gamma$ by applying $g^{\alpha\gamma}$ $= \delta^{\alpha}{}_{\beta}$

But we know that g x g = 1 so we write things out as on the right, a simple Kronecker but with appropriate position of the indices.

Now, here is what a normal LT on a contravariant vector must look like:

$x'^{\mu} = a^{\mu}{}_{\nu} x^{\nu}$ // summed indices must always be on a diagonal line

This is the normal "matrix representation" where we might write in naive simple notation that

$x'_{\mu} = A_{\mu\nu} x_{\nu}$

See "physics questions" for more notes on this subject! ( and see my long M&M chap 5 curvilinear coordinates digression notes where we always write $x'^{\mu} = T^{\mu}{}_{\nu} x^{\nu}$ and $x'_{\mu} = T_{\mu}{}^{\nu} x_{\nu}$. )

---

Now start notes on this section 2.1. It starts off defining *proper* Lorentz Transformations as those with det = +1, while space or time inversion have -1.

Next, we rewrite the Dirac equation as in 2.4 where we have defined $\gamma^{i} = \beta\alpha_{i}$ and $\gamma^{0} = \beta$. This is the obvious thing to do, we are trying to get Dirac into something resembling covariant form after all. The fact that the $\gamma$ commutators come out so nicely is encouraging! I have verified everything here including all the facts about these $\gamma^{\mu}$ guys. We have

$\partial/\partial x^{\mu} = \partial_{\mu} = \nabla_{\mu}$

$\gamma^{\mu}\partial_{\mu} = \not{\partial} = \not{\nabla}$ and $i\hbar\not{\nabla} = \not{p}$ as on page 18

Then equation 2.4 really says

$[\, i\hbar\, (\gamma^{\mu}\, \partial_{\mu}) - mc\, ]\, \psi = 0$

or

$[\, i\hbar\not{\partial} - mc\, ]\, \psi = 0$

or

$[\, i\hbar\not{\nabla} - mc\, ]\, \psi = 0$ 2.7

or

$[\not{p} - mc\, ]\, \psi = 0$ 2.8

and this have certainly gotten a lot more compact!

It certainly "looks" covariant in that $\not{p}$ seems to be a scalar in some sense.

**2.2 Proof of Covariance (18).** The statement of the problem is very clear, and I am happy to try to use the same $\gamma^{\mu}$ matrices in both the primed and unprimed frames. If we get a solution assuming that, fine. Here is how the spinors are going to transform:

$\psi'(x') = S(a)\psi(x)$

The whole thing is solved if we can find a matrix S(a) such that

$S(a)\ \gamma^{\mu}\ S^{-1}(a)\ a^{\nu}{}_{\mu} = \gamma^{\nu}$ // BD page 20 equation A

since this makes equation E on page 19 have the form $[\ i\hbar\ \partial'_{\mu} - mc]\ \psi'(x') = 0$, which is to say, this makes the Dirac equation have the same form in the primed frame as it did in the unprimed frame. This is called "form invariance" I think. If we apply S from the right and $S^{-1}$ from the left on the above, it becomes ( remember $a^{\mu}{}_{\nu}$ is just a number)

$a^{\nu}{}_{\mu}\ \gamma^{\mu} = S^{-1}(a)\ \gamma^{\nu}\ S(a)$ // BD page 20 equation 2.12

which for comparison to something below I swap sides to write as

$S^{-1}(a)\ \gamma^{\nu}\ S(a) = a^{\nu}{}_{\mu}\ \gamma^{\mu}$

**Aside**: At this point, I would like to remind myself of the rotation world where we have things like this

$R^{-1}\ \mathbf{V}\ R = R\ \mathbf{V} = \mathbf{V}'$ // compare to S thing above

where $R = R_x(\theta) = \exp(-i\theta J_x)$ for example, and $J_x$ is a generator. The thing R is a 3x3 representation of the rotation group SO(3), see sheet on TK binder. On the other hand, the object R is a 2x2 rotation matrix which operates in the $|\ ½\ ;\pm\ ½\ >$ j=1/2 "spin space" of Levitt's book. The object **V** is some "vector operator" and we are trying to see what the sandwich does to this operator.

So the point is that we have two different representations of the rotation group sort of tied together by my equation, namely, the 2D and 3D representations. This is summarized on my one-page TK binder item.

We have a similar situation going on here in BD. The 3x3 role of $R_{ij}$ is played by $a^{\nu}{}_{\mu}$ which is a 4x4 representation of the Lorentz group. The object S(a) which corresponds to my R is the representation of a LG element in the "Dirac Spinor" 4D representation. Note that it is a 4x4 representation, but is not the same as the 4x4 "vector" representation $a^{\nu}{}_{\mu}$. Just as my R is what rotates the little j=1/2 angular momentum states, this thing S(a) is going to be the thing that "rotates" the 4-spinors of the Dirac theory.

I quote from a wiki page on LG representations:

- (0,0) the Lorentz scalar representation. This representation is carried by relativistic scalar field theories.
- (1/2,0) is the left-handed Weyl spinor and (0,1/2) is the right-handed Weyl spinor representation.
- **(1/2,0) ⊕ (0,1/2) is the Dirac spinor representation.**
- **(1/2,1/2) is the vector representation. The electromagnetic vector potential lives in this rep. It is a 1-form field.**
- (1,0) is the self-dual 2-form field representation and (0,1) is the anti-self-dual 2-form field representation.

- (1,0) ⊕ (0,1) is the representation of a parity invariant 2-form field. The **electromagnetic field tensor** transforms under this representation.
- (1,1/2) ⊕ (1/2,1) is the Rarita-Schwinger field representation.
- (1,1) is the spin-2 representation of the traceless metric tensor.

So, the $a^{\nu}{}_{\mu}$ matrices are the (1/2,1/2) "vector" representation of the LG, whereas the S(a) matrices are the "Dirac Spinor" representation which is denoted (1/2,0) ⊕ (0,1/2). It happens that these are both 4x4 representations, but they are not the same.

So that is the "group theory" of what we are doing here. This is elaborated in Messiah Vol II I am happy to say, where the same metric is used for $g_{\mu\nu}$ luckily for me.

Now back to the issue of $R^{-1}\mathbf{V}\,R = R\,\mathbf{V}$ noted above. Suppose we knew the 3x3 matrices R and suppose we did *not know* the 2x2 matrices R. We could construct them starting with infinitesimals and generators, and then exponentiate to get the full R group representation elements. That is what B&D are going to do with $S^{-1}(a)\,\gamma^{\nu}\,S(a) = a^{\nu}{}_{\mu}\,\gamma^{\mu}$.

My TK binder, by the way, has some old notes on the conventions used by various authors for objects related to the LG, such as parameters and generators.

More comments. Consider again

$$S^{-1}(a)\,\gamma^{\nu}\,S(a) = a^{\nu}{}_{\mu}\,\gamma^{\mu} = \gamma'^{\nu} \qquad \psi'(x') = S(a)\psi(x) \qquad \text{BD 2.12}$$

Think of $\gamma^{\nu}$ as an *operator* in the 4D Dirac space (it after all a 4x4 matrix). S is the *operator* which -- acting in the 4D Dirac space -- makes the vector operator $\gamma^{\nu}$ become what it is in the transformed (primed) frame $\gamma'^{\nu}$.

**The infinitesimal construction method.** First, write in infinitesimal $a^{\nu}{}_{\mu}$ as in 2.13a. This is a little strange because the parameter and the generator are combined into the same object $\Delta\omega^{\nu}{}_{\mu}$ . For example, if we were doing (passive) rotations, we would say

$$R = e^{+i\theta J} \approx 1 + i\theta J \quad \text{or} \qquad R_{ij} \approx \delta_{ij} + i\theta J_{ij}$$

So we have to think of $a^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} + \Delta\omega^{\nu}{}_{\mu}$ where $\Delta\omega^{\nu}{}_{\mu} = +iuK^{\nu}{}_{\mu}$ for a boost, or $+i\theta J^{\nu}{}_{\mu}$ for a rotation, where J and K are the 4D matrices shown in my handwritten notes "summary of Lorentz Group factors" . Thus, in this strange notation, they are saying

$$\Delta\omega^{1}{}_{2} = +i\theta(J_3)^{1}{}_{2} = +i\theta\{-i\varepsilon_{312}\} = +i\theta\{-i\} = \theta = +\Delta\theta = \text{a number}$$

Of course they are using Δθ instead of θ to suggest an infinitesimal rotation at this point.
Maybe should think of it this way

$$\Delta\omega = \Sigma_{i=1,3}\,[\,+i\,u_i K_i + i\theta_i J_i\,]$$

$$\Delta\omega^{\mu}{}_{\nu} = +i\,\Sigma_{i=1,3}\,\{\,u_i\,(K_i)^{\mu}{}_{\nu} + \theta_i\,(J_i)^{\mu}{}_{\nu}\,\}$$

where this is an *arbitrary* LT with our usual 6 parameters. Later as in 2.18 you see B&D break this into a small parameter times a generator matrix. In general, we can think of $\Delta\omega^{\nu}{}_{\mu}$ as being some linear combination of the J's and K's with appropriate small parameters, and thus $\Delta\omega^{\nu}{}_{\mu}$ describes an *arbitrary* small Lorentz transformation.

Now **where does 2.13b come from?** Consider 2.3 which says

$$\delta^{\nu}{}_{\sigma} = a_{\mu}{}^{\nu} a^{\mu}{}_{\sigma} = (\delta_{\mu}{}^{\nu} + \Delta\omega_{\mu}{}^{\nu})\,(\delta^{\mu}{}_{\sigma} + \Delta\omega^{\mu}{}_{\sigma})$$

But the two deltas give the LHS so we must then have, to first order in smallness:

$$\delta_{\mu}{}^{\nu}\,\Delta\omega^{\mu}{}_{\sigma} + \Delta\omega_{\mu}{}^{\nu}\delta^{\mu}{}_{\sigma} = 0 \qquad \text{or } \Delta\omega^{\nu}{}_{\sigma} + \Delta\omega_{\sigma}{}^{\nu} = 0 \quad \text{or} \quad \Delta\omega^{\nu}{}_{\sigma} = -\Delta\omega_{\sigma}{}^{\nu}$$

We can then lower the ν index on both sides by applying $g_{\nu'\nu}$ to both sides, or by just "doing it", and this gives $\Delta\omega_{\nu\sigma} = -\Delta\omega_{\sigma\nu}$ as claimed, and this would be true in all four index combinations.

Next, **how do we justify the strange form 2.14 given for infinitesimal S**? This took me a while to figure out, see "proof of Taylor series expansion". My conclusion in those notes is that you can think about expanding a matrix F that is a function of a matrix X which is near a matrix A. Although I can write down terms of any smallness order, only the lowest order linear term is simple to write in a compact form, and we get

$$F(X) = F(A) + \Sigma_{r,s=1,N}\ (\partial F/\partial X_{rs})\ (X_{rs}-A_{rs}) = F(A) + \Sigma_{rs=1,N}\, K_{rs}\, \delta X_{rs}$$

where each $K_{rs}$ is itself a matrix just labeled by "rs". In our case here, we have A = 1 and N=4, so we have

$$F(X) = F(1) + \Sigma_{r,s=1,4}\ (\partial F/\partial X_{rs})\ (X_{rs}-\delta_{rs}) = F(1) + \Sigma_{rs=1,N}\, K_{rs}\, (X_{rs}-\delta_{rs})$$

Now we change F → S and X → a to come closer to the BD form

$$S(a) = S(1) + \Sigma_{r,s=1,4}\ (\partial S/\partial a_{rs})\ (a_{rs}-\delta_{rs}) = S(1) + \Sigma_{rs=1,N}\, K_{rs}\, (a_{rs}-\delta_{rs})$$

Next, we know that S(1) = 1, and we know $(a_{rs}-\delta_{rs}) = \Delta\omega_{rs}$ so we get

$$S(a) = 1 + \Sigma_{r,s=1,4}\ (\partial S/\partial a_{rs})\ \Delta\omega_{rs} = 1 + \Sigma_{rs=1,N}\, K_{rs}\, \Delta\omega_{rs}$$

Now we rename the 16 matrices $K_{rs} = (-i/4)\,\sigma_{rs}$ and we get

$$S(a) = 1 + \Sigma_{r,s=1,4}\ (\partial S/\partial a_{rs})\ \Delta\omega_{rs} = 1 - (i/4)\,\Sigma_{rs=1,N}\, \sigma_{rs}\, \Delta\omega_{rs}$$

where

$$-(i/4)\,\sigma_{rs} = K_{rs} = (\partial S/\partial a_{rs})|^{a=1}$$

Then finally we change to Greek letters with implied summation

$$S(a) = 1 + (\partial S/\partial a_{\mu\nu})\ \Delta\omega_{\mu\nu} = 1 - (i/4)\ \sigma_{\mu\nu}\ \Delta\omega_{\mu\nu}$$

and we can replace $\Delta\omega_{\mu\nu} = \Delta\omega^{\mu\nu}$ and we are there!

Now, since we know that $\Delta\omega_{\mu\nu}$ is antisymmetric, if we decompose $\sigma_{\mu\nu}$ into S and A parts, only the A part will survive, so we might as well use that part and then rename it to be just σ

$$\sigma_{\mu\nu}\ \Delta\omega_{\mu\nu} = [\sigma^{S}_{\mu\nu} + \sigma^{A}_{\mu\nu}]\ \Delta\omega_{\mu\nu} = \sigma^{A}_{\mu\nu}\ \Delta\omega_{\mu\nu}$$

**Comment**: I just spent 6AM to 2PM or 8 hours absorbing 4 vertical inches of B&D page 20! That is a pretty amazingly slow rate of progress, but I had to digress into other areas to "get happy". All the time I am reviewing 28 year old history back to 1980 or so. Things are a little rusty as you might expect.

**Derive bottom Equation G page 20.** All we need to do here is insert the three through-first-order expressions for S, $S^{-1}$ and a into 2.12. The leading terms on both sides cancel, and then we set the first order terms to 0 and this gives the result shown on the bottom of the page.

**Derive page 21 equation 2.15.** Write the LHS of page 20 equation G as shown in pencil in the margin, then both sides have $\Delta\omega^{\alpha\beta}$ and we set the other parts equal. Before we cancel off the $\Delta\omega^{\alpha\beta}$, we use the fact that $\Delta\omega^{\alpha\beta}$ is A (antisymmetric) so write it's mating factor as A + S and keep only the A part as usual, and we get the 2.15 result.

**Show that 2.16 satisfies 2.15** . At first look this is totally non-obvious. You start off with this:

$$RHS = (i/2)\ [\gamma^{\mu}, [\ \gamma^{\alpha}, \gamma^{\beta}]\ ]$$

where I just raised the indices on 2.16 and 2.15. But this is a sum of four terms each completely general products of three gammas, how do you reduce this to the LHS which has only one gamma and a metric tensor. This right now is very non-obvious to me, but I have the thing checked from last time I was here. The Jacobi identity does not seem to help. You somehow have to use the $\{\ \gamma^{\alpha}, \gamma^{\beta}\ \} = 2g^{\alpha\beta}$ fact to remove pairs of gammas.

For **compact notation** (this is my own deal), just write the index and omit the γ so we have

$$[\nu,[\alpha,\beta]] = [\nu, \alpha\beta - \beta\alpha\ ] = \nu(\alpha\beta - \beta\alpha) - (\alpha\beta - \beta\alpha)\nu = \nu\alpha\beta - \nu\beta\alpha - \alpha\beta\nu + \beta\alpha\nu = 1\ \ 2\ \ 3\ \ 4$$

$$= (\ \nu\alpha\beta - \alpha\beta\nu\ ) - (\ \nu\beta\alpha - \beta\alpha\nu\ )\quad 1\ \ 3\ \ 2\ \ 4$$

Each time we move a letter one position, we pick up a minus sign and some g tensor stuff. Suppose we take the third term and slide ν to the left two positions. We end up with $-\nu\alpha\beta$ which cancels the first term, and we pick up two g things. Then suppose we take the last term and move ν two positions to the left. This will cancel the second term, and we are left only with g stuff, so that is the solution. Now go do it in more detail:

the rule is $\alpha\beta + \alpha\beta = 2g_{\alpha\beta}$ so $\alpha\beta = -\alpha\beta + 2g_{\alpha\beta}$ or $\alpha\beta = -\alpha\beta + 2ab$

where ab is my shorthand for $g_{\alpha\beta}$. So, start with term 3. We have

$$- \alpha\beta\nu = -\alpha(-\nu\beta + 2b\nu) = +\alpha\nu\beta - 2\alpha(bv)$$

Now move the ν to the left in ανβ to get

$$\alpha\nu\beta = (-\nu\alpha + 2va)\beta = -\nu\alpha\beta + 2\beta(va)$$

Thus, our third term becomes

$$- \alpha\beta\nu = \alpha\nu\beta - 2\alpha(bv) = -\nu\alpha\beta + 2\beta(va) - 2\alpha(bv)$$

And we can then say

$$1st + 3rd\ term = \nu\alpha\beta - \nu\alpha\beta + 2\beta(va) - 2\alpha(bv) = + 2\beta(va) - 2\alpha(bv)$$

We can handle the 2nd + 4th term by taking the previous result, swapping α and β (and a and b) and adding an overall minus sign, as shown on second isolated line about 6 inches up this page. So

$$2nd + 4th\ term = -\{ 2\alpha(vB) - 2\beta(av)\}$$

Therefore, the sum of all four terms is this:

$$1+2+3+4 = + 2\beta(va) - 2\alpha(bv) - 2\alpha(vb) + 2\beta(av) = 4\beta\, va - 4\alpha\, bv = 4(\beta\, va - \alpha\, bv)$$

Thus we have shown that

$$[\nu,[\alpha,\beta]] = 4(\beta\, va - \alpha\, bv)$$

which we now decompress to get

$$[\gamma^\nu,[\gamma^\alpha,\gamma^\beta]] = 4(\gamma^\beta\, g^{\nu\alpha} - \gamma^\alpha\, g^{\nu\beta})$$

Therefore we have

$$[\gamma^\nu, \sigma^{\alpha\beta}] = (i/2)\,[\gamma^\nu,[\gamma^\alpha,\gamma^\beta]] = 2i\,(\gamma^\beta\, g^{\nu\alpha} - \gamma^\alpha\, g^{\nu\beta})$$

We can now lower the αβ indices if we want to get

$$[\gamma^\nu, \sigma_{\alpha\beta}] = (i/2)\,[\gamma^\nu,[\gamma_\alpha,\gamma_\beta]] = 2i\,(\gamma_\beta\, g^\nu{}_\alpha - \gamma_\alpha\, g^\nu{}_\beta) \quad \text{// verifying } \sigma \text{ makes 2.15}$$

and this is the way things appear in 2.15, **QED** and add this to the gamma page please! So this is an important gamma matrix reduction formula reducing triple products to singles!

So at this point, we have found an explicit form for the infinitesimal S.

Now consider a *passive* x-axis boost by u. Look at TK "summary of LG facts" where we have

$B_x(\omega) = \exp(+i\omega K_1)$ where $K_1$ = as shown in those notes. Note that $iK_1$ = BD 2.19 for I. So

$B_x(\omega) = \exp(+\omega i K_1) = \exp(+\omega\ I)$ = my matrix with four + signs.

So looking at my notes for $B_x(u)$, we get exactly 2.20 of BD. I can skip the details of getting inf to finite because I know what the answer is already. So we are now mid page 22.

**But** let's look at this interesting method anyway. He considers a set of N small transformations of the infinitesimal form and sets $\Delta\omega = \omega/N$ as shown. Why does this become exponential in the limit? The implicit theorem is this:

$$\lim_{N\to\infty} (1 + [\omega/N]\ x)^N = e^{\omega x}$$

The way to prove this is to expand the LHS using the binomial expansion, take the large N limit, and this becomes the expansion of $e^{\omega x}$. For example

$$(1 + [\omega/N]\ x)^N = 1 + N\ (\omega x/N) + N(N-1)/2\ (\omega x/N)^2 + N(N-1)(N-3)/3!\ (\omega x/N)^3 + ....$$

$$\rightarrow\ 1 + (\omega x) + (\omega x)^2/2! + (\omega x)^3/3! + ...\ \text{QED}$$

The last step comes from this fact

$$ch(\omega I) = 1 + (\omega I)^2/2! + (\omega I)^4/4! + .... = 1 + \{ ch(\omega) - 1 \}\ I^2$$

and when you look at matrices I and $I^2$ on page 21 you can fill everything in to get 2.20 which is the correct form of a passive boost.

Suppose we took a particle at rest p = (m,0,0,0) and boost it by v as shown in 2.20. We know that we get that p' = ($\gamma$mc, $\gamma$mv,0,0). We know then that $\gamma$ = cosh($\omega$) and $\gamma\beta$ = sinh($\omega$) and tanh($\omega$) = $\beta$. So this confirms equation E on page 22.

Now let's back up and be more general. Looking back at 2.13 we had

$$a = 1 + \Delta\omega + .... = \exp(\omega\ (I_n)) \quad \text{where } \Delta\omega^{\mu\nu} = \Delta\omega\ (I_n)^{\mu\nu} \quad \Delta\omega = \omega/N$$

Similarly, we are going to have

$$S = 1 - (i/4)\sigma_{\mu\nu}\Delta\omega^{\mu\nu} + .... = \exp[\ -(i/4)\ \omega\ \sigma_{\mu\nu}\ (I_n)^{\mu\nu}\ ]$$

as shown in 2.22.

Now look again at 2.19 for $I^\nu{}_\mu$. What is $I^{\nu\mu}$ ?

$$I^{\nu\mu} = g^{\mu\alpha}\ I^\nu{}_\alpha$$

But $I^\nu{}_\alpha = -1\ (\ \delta_{\nu,0}\delta_{\alpha,1} + \delta_{\nu,1}\delta_{\alpha,0})$

So

$$\Gamma^{\nu\mu} = g^{\mu\alpha}\,\Gamma^{\nu}{}_{\alpha} = -1\; g^{\mu\alpha}\,(\,\delta_{\nu,0}\delta_{\alpha,1} + \delta_{\nu,1}\delta_{\alpha,0}) = -1\,(\delta_{\nu,0}\,g^{\mu 1} + \delta_{\nu,1}\,g^{\mu 0})$$

$$= -1\,(\,-\,\delta_{\nu,0}\,\delta_{\mu,1} + \delta_{\nu,1}\,\delta_{\mu,0})$$

$$= (\,\delta_{\nu,0}\,\delta_{\mu,1} - \delta_{\nu,1}\,\delta_{\mu,0})$$

Therefore

$$\sigma_{\nu\mu}\,(I_n)^{\nu\mu} = \sigma_{\nu\mu}\,(\,\delta_{\nu,0}\,\delta_{\mu,1} - \delta_{\nu,1}\,\delta_{\mu,0}) = \sigma_{01} - \sigma_{10} = 2\,\sigma_{01}$$

and then in this case we have

$$S = \exp[\,-\,(i/4)\,\omega\,\sigma_{\mu\nu}\,(I_n)^{\mu\nu}\,] = \exp[\,-\,(i/4)\,\omega\;\;2\,\sigma_{01}\,] = \exp[\,-\,(i/2)\,\omega\,\sigma_{01}]$$

**just as claimed in 2.23**. So this tells you what a passive x boost does to a 4-spinor!

Now let's repeat this last exercise when $I = iJ_3$ instead of $iK_1$. My hand written notes say that

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}$$

so looking at my rotation TK summary page, I know that

$$J_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad I = iJ_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \Gamma^{\nu}{}_{\mu} = \text{like 2.19.}$$

We can see that

$$\Gamma^{\nu}{}_{\alpha} = 1\,(\,\delta_{\nu,1}\delta_{\alpha,2} - \delta_{\nu,2}\delta_{\alpha,1})$$

$$\Gamma^{\nu\mu} = g^{\mu\alpha}\,\Gamma^{\nu}{}_{\alpha} = 1\; g^{\mu\alpha}\,(\,\delta_{\nu,1}\delta_{\alpha,2} - \delta_{\nu,2}\delta_{\alpha,1})$$

$$= -1\;\;(\,\delta_{\nu,1}\delta_{\mu,2} - \delta_{\nu,2}\delta_{\mu,1}) = (\,\delta_{\nu,2}\delta_{\mu,1} - \delta_{\nu,1}\delta_{\mu,2})$$

$$\sigma_{\nu\mu}\,(I_n)^{\nu\mu} = \sigma_{\nu\mu}\,(\,\delta_{\nu,2}\delta_{\mu,1} - \delta_{\nu,1}\delta_{\mu,2}) = \sigma_{21} - \sigma_{12} = -\,2\,\sigma_{12}$$

$S = \exp[ - (i/4)\, \omega\, \sigma_{\mu\nu}\, (I_n)^{\mu\nu} ] = \exp[ + (i/4)\, \omega\ \ 2\, \sigma_{12} ] = \exp[ + (i/2)\, \omega\, \sigma_{12}]$

and this agrees with 2.24. Note that $\sigma_{12} = \sigma^{12}$ and they have $\sigma^{12}$, fine.

Now compute $\sigma^{12} = (i/2)\, [\, \gamma^1, \gamma^2] = i\, \gamma^1\gamma^2 = \Sigma^3$ from my hand notes and this is as shown below 2.24.

I agree that the $\sigma_{ij}$ are hermitian and that the rotations are unitary so $S^\dagger = S^{-1}$ . And I agree that the $\sigma_{0j}$ are antihermitian, so we get $S^\dagger = S$ instead.

They now make the claim that

$S^{-1} = \gamma_0 S^\dagger \gamma_0$

is true for both rotations and boosts. Here is a proof:

**Lemma:** claim that $\{\gamma_o, \sigma_{0j}\} = 0$.

Proof: $\{\gamma_o, \sigma_{0j}\} \stackrel{\bullet}{=} \{\, \gamma_o\, , [\gamma_0, \gamma_j]\, \} = \{\, 0, 0j - j0\} = 0(0j - j0) + (0j - j0)0$

$= 00j - 0j0 + 0j0 \ - j00 \ = \ j - 0j0 + 0j0 \ - j = 0$ QED

This tells us that

$\gamma_0\, \sigma_{0j}\gamma_0 \ = -\, \sigma_{0j}$

Therefore, when we take a boost $S_L$ and sandwich $\gamma_0$ around it, we change the sign in the exponent, and that is what we need to show that

$S^{-1} = e^{+} = \gamma_0 S^\dagger \gamma_0 = \gamma_0 S \gamma_0 = \gamma_0 e^{-} \gamma_0 = e^{+}$ for boosts

For rotations, we know $\sigma_{ij}$ commutes with $\gamma_o$ so $\gamma_0\, \sigma_{ij}\gamma_0 \ = +\, \sigma_{ij}$ so the sandwich does nothing at all, and we have $S^{-1} \ = \ \gamma_0 S^\dagger \gamma_0 \ = S^\dagger$ for rotations.

**Therefore, as claimed, 2.26 is valid** for both boosts and rotations.

BD then show **easily and clearly that the 4-current $j^\mu$ transforms as a 4-vector** as we expected, and we use our S rule 2.12 to show this. We already know that $\partial_\mu j^\mu = 0$ and this is then seen as a scalar condition.

Anticipating that $\psi^\dagger\gamma_0$ shows up a lot, we call it $\bar{\psi} = \psi^\dagger \gamma_0$. One example is that it shows up in the current so we really have

$j^\mu \ = c\, \bar{\psi}\, \gamma^\mu \psi$

**Review what we learned in this section!**

For the normal vector representation of the Lorentz Group (LG), we denote the usual (passive) transformation matrices by a and elements by $a^{\mu}{}_{\nu}$ with "the usual tilt" for a normal matrix. The matrices a are representations of elements of the LG. They act on 4-vectors (not Dirac 4-spinors). The "vector" representation of the LG has the designation **(1/2,1/2)**. Remember that the Lorentz Group has two casimirs the way the rotation group has one casimir $J^2$ and the group designation lists the values of the casimirs.

The 4-spinors $\psi$ do not -- although they have four elements -- transform as 4-vectors. Instead, they transform according to a different representation of the LG known as the "Dirac Spinor" representation which has the group theoretic designation **(1/2,0)** $\oplus$ **(0,1/2)**. It would appear that this is a "reducible" representation of the LG.

As we change reference frame, we know that a Lorentz Group vector object transforms as $x'^{\mu} = a^{\mu}{}_{\nu}x^{\nu}$ . We could write this as x' = ax. Think of x as if it were a ket, and we are saying $x'^{\mu} = a^{\mu}{}_{\nu}x^{\nu}$ in the sense of acting on a vector in Hilbert space which vector is labeled by the index x. If we wanted to deal with a "vector operator" $V^{\mu}$ in this space, we would say $A^{-1}V\,A = a\,V = V'$. Normally we don't do this, however. For example, the current $j^{\mu}$ at least in this book is not an operator, it is a 4-vector ket thing and it transforms the same way that $x^{\mu}$ does.

The *generators* of the vector transformations are written in this section as $\Delta\omega^{\nu}{}_{\mu} = \delta\omega\, I^{\nu}{}_{\mu}$ so really the $I^{\nu}{}_{\mu}$ are the generator matrices which I like to call J and K in my hand written notes, and they are also sometimes written in $M_{\mu\nu}$ form but we have avoided that so far. BD don't say much about the generators $I_n$, but we can think of n = 1,2,3,4,5,6 if we like.

As we change reference frame, we know that a Lorentz Group Dirac Spinor object transforms as $\psi'(x') = S(a)\psi(x)$ or write this as $\psi' = S\,\psi$. This is in the "ket sense" mentioned in the previous paragraph. If we had a "Dirac Spinor operator" D, it would transform as $S^{-1}D\,S = a\,D = D'$ . In this chapter, we have discovered that the four matrices $\gamma^{\mu}$ do in fact form a Dirac Spinor Operator, and these matrices transform in exactly this way as shown in equation 2.12.

The *generators* of the Dirac Spinor transformations are written as $G_n = (-i/4)\,\sigma_{\mu\nu}(I_n)^{\mu\nu}$ again for n = 1,2,3,4,5,6. The $(I_n)^{\mu\nu}$ are numbers -- elements of the *Vector* generator matrices. The $\sigma_{\mu\nu}$ are matrices for each choice of $\mu\nu$. So of course $G_n$ are 4x4 matrices which act as generators in this Dirac Spinor space! BD never write down the 4x4 $G_n$ matrices explicitly, but instead they look at special cases. For example, when we have a passive x boost, $G_{B1} = (-i/4)\,2\,\sigma_{01}$, and when we have a passive z rotation, $G_{R3} = -\,(-i/4)\,2\,\sigma_{12}$. From these examples, I think we conclude that

$$G^{boost}{}_{i} = (-i/4)2\sigma_{0i}$$
$$G^{rot}{}_{i} = -\,2\,(-i/4)\,\varepsilon_{ijk}\sigma_{jk}$$

so in essence that we can regard $\sigma_{\mu\nu}$ as *being* the Dirac Spinor generators. Since antisymmetric, there are only (16-4)/2 = 6 different matrices, and they are the generators as shown (times constant).

It happens that $\sigma_{\mu\nu} = (i/2)\,[\gamma_{\mu}, \gamma_{\nu}]$ so we could say that the Dirac Spinor generators are given as a commutator of some Dirac Spinor operators!

The theme of this section is that we can show the Dirac equation is covariant if we can show it has the same form in primed and unprimed frames, AND if we can show how someone in one frame can compute what the other frame's observer is seeing, $\psi'(x') = S\,\psi(x)$, but of course this last item is just saying that we want to know all about the details of things, not just the fact that something is covariant in form. We start by showing that we can show the form of the Dirac equation is covariant if we can find S such that 2.12 is true. Now I suppose we might have found that no such S exists, and then our covariance proof would fail. But this is not what we found.

When we form the following grouping $k^\mu = \bar{\psi}\,\gamma^\mu\,\psi$ where $\bar{\psi} = \psi^\dagger\gamma_0$ , we have learned that although $\psi$ is a ket of the Dirac Spinor Representation, and although $\gamma^\mu$ transforms as a Dirac Spinor Operator, the combination $k^\mu$ transforms as a Vector object! This is shown very precisely in 2.27 ! The reason of course is that the operator and ket transformations "cancel each other". Here is a fast way to see it

$$k^{\mu\,\prime} = \langle\bar{\psi}'\,|\,\gamma^\mu\,|\,\psi'\rangle = \langle\bar{\psi}\,|\,S^{-1}\gamma^\mu\,S\,|\,\psi\rangle = a^\mu{}_\nu\langle\bar{\psi}\,|\,\gamma^\mu\,|\,\psi\rangle = a^\mu{}_\nu\,k^\nu = \langle\bar{\psi}\,|\,\gamma^{\mu\prime}|\,\psi\rangle$$

**2.3 Space Reflection (24).** What S satisfies 2.12 for an "a" (ie, $a^\mu{}_\nu$) which just inverts the spatial coordinates? $S = \gamma^0 \equiv P$ does just fine, to which you can add a phase if you want to get 2.32. $\psi$ is not an eigenstate of this P operator. However, if you look at the four rest-state solutions on page 10, the positive energy ones have positive parity, and the negative energy states have negative parity if we choose $P = +\gamma_0$. This is just because $P = \gamma^0 = \mathrm{diag}(1,-1)$ where 1 is a 2x2. Any linear combination of the first two states has positive parity, so this would include any non-rel-limit solution of the Dirac equation.

Time Reflection? We have to wait until chapter 6 to ponder this, not sure why we could not just do it right here.

**2.4 Bilinear Covariants (25)**. In this section we examine 16 independent 4x4 matrices which span the Dirac spinor space. They are put into 5 groups where the superscript indicates how the things in that group transform when you sandwich them between $\bar{\psi}$ and $\psi$. S = scalar, P = pseudoscalar, A = axial vector (that is to say, a pseudovector), V = vector (ie, a polar vector), T = tensor (of rank 2). Again, we are not saying that the 4x4 matrix $\Gamma^\nu{}_\mu = \gamma_\mu$ *itself* transforms as a vector, but that $\bar{\psi}\,\gamma_\mu\,\psi$ does. BD provide a list of properties on all these matrices (which can all be expressed in terms of the four $\gamma^\mu$ and $\gamma^0$. Note that $\gamma_5$ is a certain combination of the others and $\gamma_5$ = offdiag(1,1) so it swaps the upper and lower component pairs of whatever Dirac spinor thing it acts on. ) Given the list of properties, they prove easily that the 16 matrices $\Gamma$ really are linearly independent and thus span the space. Any combination $\bar{\psi}$ (stuff) $\psi$ is referred to as being "bilinear" because the $\psi$ appears twice. The combinations on page 26 can each be identified with some transformation rule under the Lorentz group "a". Since each of these bilinear forms transforms in some well-defined group manner, they are called "covariants". The factor det(a) is used to provide a minus sign for a parity transformation and a plus sign for all proper Lorentz transformations.

## Chapter 3: Solutions of the Dirac Equation for a Free Particle (27)

**3.1 Plane Wave Solutions (28).**

**Comment**: We learned in previous chapters how Dirac stumbled upon the "Dirac equation". The problem was to find a Hamiltonian H to put into the SE which says $H\psi = E\psi$, both operators. The H had to respect the relativistic formula for energy E such that

$$E^2\psi = H^2\psi = (\mathbf{p}^2c^2 + m^2c^4)\psi = ([-i\hbar\nabla]^2c^2 + m^2c^4)\psi$$

and H had to result in a conserved probability current $j_\mu$ with a positive definite $\rho$. "Early attempts" had various failure modes. By going with a 4 dimensional matrix approach, Dirac found a theory that met all the requirements. We have since come to understand that $\psi$ "belongs" to the **(1/2,0) ⊕ (0,1/2)** representation of the Lorentz group.

As before, we first deal with technical details, then resume notes on this section.

**Derive 3.5 .** First, we have $S = \exp(-i\omega\sigma_{01}/2)$ for an x boost which we confirmed in 2.23, but I thought that was passive and now this one is active. Well, keep with the above sign for the moment. We find that

$\sigma_{01} = i\gamma_0\gamma_1$ // from our sheet
$= -i\,\gamma^0\gamma^1 = -i\alpha_1$ $\Rightarrow\ -i\sigma_{01} = (-i)(-i\alpha_1) = -\alpha_1$

so we want $S = \exp(-i\omega\sigma_{01}/2) = \exp(-\alpha_1\omega/2)$. So now we have to go off and look at powers of $\alpha_1$.

$\alpha_1 = \gamma^0\gamma^1$
$\alpha_1^2 = \gamma^0\gamma^1\gamma^0\gamma^1 = -\gamma^0\gamma^0\gamma^1\gamma^1 = -(+1)(-1) = 1$
$\alpha_1^3 = \alpha_1$
$\alpha_1^4 = 1$

Thus we have

$$\exp(-\alpha_1\omega/2) = \Sigma_{n=0,\infty}\ (-\omega/2)^2/n!\ \ \alpha_1^n$$

$$\text{sum even terms} = \Sigma_{n=0,2,\infty}\ (-\omega/2)^2/n! * 1 = \text{ch}(-\omega/2) * 1$$

$$\text{sum odd terms} = \Sigma_{n=1,3,\infty}\ (-\omega/2)^2/n! * \alpha_1 = \text{sh}(-\omega/2) * \alpha_1$$

Therefore

$$\exp(-\alpha_1\omega/2) = \text{ch}(\omega/2) * 1 - \text{sh}(\omega/2) * \alpha_1$$ // agrees with page 29 equation A

But the matrix $\alpha_1$ is this one:

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \exp(-\alpha_1\omega/2) = \begin{pmatrix} \text{ch}(\omega/2) & 0 & 0 & -\text{sh}(\omega/2) \\ 0 & \text{ch}(\omega/2) & -\text{sh}(\omega/2) & 0 \\ 0 & -\text{sh}(\omega/2) & \text{ch}(\omega/2) & 0 \\ -\text{sh}(\omega/2) & 0 & 0 & \text{ch}(\omega/2) \end{pmatrix}$$

and this is exactly equation 3.5 where the ch(ω/2) has been factored out.

BD now want to continue with velocity boost in an arbitrary direction, and I accept page 29 C as the right matrix because 2.19 showed us what it was for a pure x where we had -1 where -cosα sits. I am fine for the first equality in 3.7, and now I have to do this:

**Compute** exp[ -(ω/2) $\boldsymbol{\alpha}\bullet\hat{\mathbf{v}}$ ] **as in 3.7:**

Well, things are pretty simple because

$$\boldsymbol{\alpha}\bullet\hat{\mathbf{v}} = \begin{pmatrix} 0 & \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} \\ \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} & 0 \end{pmatrix} \qquad \text{and then } (\boldsymbol{\alpha}\bullet\hat{\mathbf{v}})^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{since } (\boldsymbol{\sigma}\bullet\hat{\mathbf{v}})^2 = 1$$

sum even terms = $\Sigma_{n=0,2,\infty}\ (-\omega/2)^2/n! * 1 = \mathrm{ch}(-\omega/2) * 1$

sum odd terms = $\Sigma_{n=1,3,\infty}\ (-\omega/2)^2/n! * \alpha_1 = \mathrm{sh}(-\omega/2) * \boldsymbol{\alpha}\bullet\hat{\mathbf{v}}$

so that

$$\exp[\ -(\omega/2)\ \boldsymbol{\alpha}\bullet\hat{\mathbf{v}}\ ] = \mathrm{ch}(\omega/2)*1 - \mathrm{sh}(\omega/2) * \boldsymbol{\alpha}\bullet\hat{\mathbf{v}}$$

write this now as

$$\mathrm{ch}(\omega/2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \mathrm{sh}(\omega/2)\begin{pmatrix} 0 & \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} \\ \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} & 0 \end{pmatrix} = \mathrm{ch}(\omega/2)\ \{\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \mathrm{th}(\omega/2)\begin{pmatrix} 0 & \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} \\ \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} & 0 \end{pmatrix}\}$$

Now we know we can write

$$\boldsymbol{\sigma}\bullet\hat{\mathbf{v}} = \boldsymbol{\sigma}\bullet\mathbf{v}/\ v = (1/v)\begin{pmatrix} v_3 & v_- \\ v_+ & -v_3 \end{pmatrix} \qquad \text{where } v_\pm = v_x \pm i v_y$$

$$= (1/p)\begin{pmatrix} p_3 & p_- \\ p_+ & -p_3 \end{pmatrix} \quad \text{where } p_\pm = p_x \pm i p_y$$

Now from previous page we have T

$$-\mathrm{th}(\omega/2) / p = c/(E+mc^2)$$

so we put this factor $c/(E+mc^2)$ into all the relevant terms in $-\mathrm{th}(\omega/2)\begin{pmatrix} 0 & \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} \\ \boldsymbol{\sigma}\bullet\hat{\mathbf{v}} & 0 \end{pmatrix}$. Then the outside factor ch(ω/2) is as shown in 3.6 and we have now exactly produced 3.7 ! We can write this in a compact notation this way:

$$3.7 = \begin{pmatrix} \text{ch}(\omega/2) & -\text{th}(\omega/2)\boldsymbol{\sigma}\bullet\hat{\mathbf{v}} \\ -\text{th}(\omega/2)\boldsymbol{\sigma}\bullet\hat{\mathbf{v}} & \text{ch}(\omega/2) \end{pmatrix}$$

**Review**: We start with $\psi(x)$ for particle at rest. We boost our observation frame in $-\hat{\mathbf{p}}$ direction so from frame F' we see electron with momentum $+\mathbf{p}$. I prefer just doing an "active boost" on the particle to get it up to this speed, but fine. In the rest frame, F, we already solved the Dirac equation on page 10 [ also page 28 where $\varepsilon_r = +1$ for r = 1,2] and found the various "unit spinors" times their phases. This phase sign is minus for r = 1,2 and these were called "the positive energy solutions".

So the idea here is this: (1) we already solved the Dirac equation for $\mathbf{p} = 0$ and we got unit spinors times covariant phaser. (2) We know what S is for an arbitrary-direction boost -- it is $S = \exp[ -(i/4)\,\omega\,\sigma_{\mu\nu}(I_n)^{\mu\nu}]$ where we use $I_n$ as appropriate for an arbitrary-direction boost as shown in page 29 C. (3) We compute this S as shown in 3.7. (4) This matrix S then converts the rest solutions to non-rest solutions. The unit spinors each become a column in this matrix S. (5) Thus, we now have the solution of the Dirac equation for a free particle of arbitrary momentum $\mathbf{p}$. I guess this was easier than trying to solve the equation directly for a particle doing momentum $\mathbf{p}$.

Something to note: Our four starting unit spinors now have the appearance as in 3.7 page 30. Recall now that

$$\mathbf{p} = \gamma\boldsymbol{\beta}mc \qquad E = \gamma mc^2 \text{ so } \quad c\mathbf{p}/E = \boldsymbol{\beta}$$

At very high energy, the non-unity factors in 4x4 matrix 3.7 have these limits

$p_z$ * factor $\rightarrow \cos\theta$ // that is, $p_z c/(E+mc^2) \rightarrow p\cos\theta c/E = \cos\theta\,\beta = \cos\theta$

$p^{\pm}$ * factor $\rightarrow \sin\theta\cos\varphi \pm i\sin\theta\sin\varphi = \sin\theta\, e^{\pm i\varphi}$

so that non-unity spinor entries are now "as large as" the unit entries, and strange things perhaps will happen. BD did not take this limit.

As we did our boost, our exponential phase *started* at $-i\varepsilon_r mc^2 t/\hbar = -i\varepsilon_r (mc)(ct)/\hbar$. In this starting frame F, we know that $x^\mu = x^\mu = (x^0, \mathbf{r}) = (ct, \mathbf{r})$ and $p^\mu = (mc,0)$, so phase is $-i\varepsilon_r x^\mu p_\mu/\hbar$ at least in the rest frame. Then this is the form in the boosted frame, because the phase has to be a world scalar! If it were something else, I think you would have a big mess.

So we now break off the spinors from the phase factor and call them $w^r(p)$.

**Derive 3.9a.** This equation says that you get 0 when you apply $(\not{p} - \varepsilon_r mc)$ to any of our four solution spinors $w^r(p)$, you get 0. Here p is just $p^\mu$, not an operator. So this equation says that our solutions solve the Dirac equation "in momentum space". This sort of verifies that we found the correct solutions.

From earlier, we had this Dirac equation:

$$[\, i\hbar\not{\partial} - mc\,]\,\psi = 0 \qquad [\, i\hbar\nabla - mc\,]\,\psi = 0 \qquad [\hat{p} - mc\,]\,\psi = 0$$

all operators. Now if we have $\exp( -i\varepsilon_r x^\mu p_\mu/\hbar)$, then

$$i\hbar\not{\partial}\exp( -i\varepsilon_r x^\mu p_\mu/\hbar) = i\hbar\gamma^\mu\partial_\mu \exp( -i\varepsilon_r x^\mu p_\mu/\hbar) = i\hbar\gamma^\mu \partial/\partial x^\mu \exp( -i\varepsilon_r x^\mu p_\mu/\hbar)$$

$= i\hbar\,\gamma^\mu(-i\varepsilon_r p_\mu/\hbar)\exp(-i\varepsilon_r x^\mu p_\mu/\hbar) = \varepsilon_r\gamma^\mu p_\mu \exp(-i\varepsilon_r x^\mu p_\mu/\hbar) = \varepsilon_r \not{p}\exp(-i\varepsilon_r x^\mu p_\mu/\hbar)$

Thus, the operator $\hat{\not{p}}$ is replaced by non-operator $\varepsilon_r \not{p}$ for a plane wave solution. So the Dirac equation can be written:

$[\varepsilon_r\not{p} - mc]\, w^r(p) = 0$ which is basically 3.9a

Here is how you would transpose this thing (note that $p_\mu$ is real)

$[\not{p}-\varepsilon_r mc]\, w^r(p) = 0 \Rightarrow [\gamma^\mu p_\mu - \varepsilon_r mc]\, w^r(p) = 0 \Rightarrow \quad w^r(p)^\dagger [\gamma^\mu p_\mu - \varepsilon_r mc]^\dagger = 0$

$\Rightarrow w^r(p)^\dagger [\gamma^{\mu\dagger} p_\mu - \varepsilon_r mc] = 0$

Apply $\gamma^0$ from the right to get

$\Rightarrow w^r(p)^\dagger [\gamma^{\mu\dagger} p_\mu - \varepsilon_r mc]\,\gamma^0 = 0 \Rightarrow w^r(p)^\dagger \gamma^0\gamma^0 [\gamma^{\mu\dagger} p_\mu - \varepsilon_r mc]\,\gamma^0 = 0$

$\Rightarrow \bar{w}^r(p)\,\gamma^0 [\gamma^{\mu\dagger} p_\mu - \varepsilon_r mc]\,\gamma^0 = 0 \Rightarrow \bar{w}^r(p)[\gamma^\mu p_\mu - \varepsilon_r mc] = 0$ QED.

**Derive 3.9b.** The claim here is that $\bar{w}^r(p)\, w^{r'}(p) = \varepsilon_r \delta_{r,r'}$. That is to say, the column vectors in 3.7 are orthogonal and almost orthonormal except for the $\varepsilon_r$ sign thing. We know the thing above is a scalar so we can evaluate it in the rest frame. We have

$\bar{w}^r(p)\, w^{r'}(p) = w^{\dagger r}(p)\,\gamma^0\, w^{r'}(p) = \langle \bar{w}^r(p) \mid w^{r'}(p)\rangle$

The $\gamma^0$ matrix has the form diag(1,-1) and this is then where the $\varepsilon_r$ comes from QED. So I like to write this result as

$\langle \bar{w}^r(p) \mid w^{r'}(p)\rangle = \varepsilon_r \delta_{r,r'}$ // orthogonality in the 4x4 Dirac space

**Derive 3.9c.** Claiming that

$\Sigma_{r=1,4}\, \varepsilon_r\, w^r{}_\alpha(p)\, \bar{w}^r(p)_\beta = \delta_{\alpha,\beta}$

My inclination is to write this as follows:

$\Sigma_{r=1,4}\, \varepsilon_r \langle\alpha| w^r(p)\rangle\langle \bar{w}^r(p)|\beta\rangle = \delta_{\alpha,\beta}$ // completeness in the 4x4 Dirac space

where you can think of $|\beta\rangle$ as being a unit column vector such as $|\beta\rangle = (1,0,0,0)$ -- you put in whatever unit vectors you want for $\alpha$ and $\beta$. Then the operator form would be

$\Sigma_{r=1,4}\, \varepsilon_r \mid w^r(p)\rangle\langle \bar{w}^r(p) \mid = 1$

which has the general form of a "completeness relation". To "prove this", I would first apply $|\, w^{r'}(p) >$ on the right to get

$$\Sigma_{r=1,4}\ \varepsilon_r\ |\, w^r(p)\rangle\langle\, \bar{w}^r(p)\, |\, w^{r'}(p)\, \rangle\ =\ |\, w^{r'}(p)\, \rangle$$

Now use the orthogonality 3.9b shown above to get

$$\Sigma_{r=1,4}\ \varepsilon_r\ |\, w^r(p)\rangle\ \varepsilon_r\ \delta_{r,r'}\ =\ |\, w^{r'}(p)\, \rangle$$

But the LHS is just $|\, w^{r'}(p)\rangle$ .

**Note Added:** Here is another way to think of the above completeness relation. We can write

$$\Sigma_{r=1,2}\ (+1)\ |\, w^r(p)\rangle\langle\, \bar{w}^r(p)\, |\ =\ 1^+$$
$$\Sigma_{r=3,4}\ (-1)\ |\, w^r(p)\rangle\langle\, \bar{w}^r(p)\, |\ =\ 1^-$$

$$1 = 1^+ + 1^-$$

where we break things down into the positive and negative energy solution spaces. If you are dealing only with positive energy spinors, we know that $w^r(p)$ with r = 1 and 2 form a complete basis, so only these need be included in the operator completeness statement above. You formally prove this by considering arbitrary linear combinations of positive energy spinors, etc etc.

Now notice this fact

$1^+\ |\, w^3(p)\rangle\ = 0$ because $\langle\, \bar{w}^r(p)\, |\, w^3(p)\, \rangle = 0$ for r = 1 and 2.

$1^+\ |\, w^1(p)\rangle\ =\ |\, w^1(p)\rangle$ because that is what $1^+$ means! And just do it.

We may therefore conclude that $1^+ = \Lambda^+$ , the positive energy projection operator (to be introduced on page 33). So although these forms don't *look* very similar, they must be the same. Thus we have added our own little fact here:

$$\Sigma_{r=1,4}\ \varepsilon_r\ w^r{}_\alpha(p)\ \bar{w}^r(p)_\beta\ = \delta_{\alpha,\beta}$$

$$\Lambda^+(p) = (\not{p}+mc)/2mc\ \ = \Sigma_{r=1,2}\ |\, w^r(p)\rangle\langle\, \bar{w}^r(p)\, |\ =\ \Sigma_{r=1,2}\ w^r(p)\ \bar{w}^r(p)\ = 1^+ = +\Sigma_s\ u(p,s)\ \bar{u}(p,s)$$

$$\Lambda^-(p) = (-\not{p}+mc)/2mc\ = -\ \Sigma_{r=3,4}\ |\, w^r(p)\rangle\langle\, \bar{w}^r(p)\, |\ = -\ \Sigma_{r=3,4}\ w^r(p)\ \bar{w}^r(p)\ =\ 1^- = -\Sigma_s\ v(p,s)\ \bar{v}(p,s)$$

Notice the minus sign in the last line, and notice that everything is a 4x4 matrix. These results are used on page 95 to obtain 6.48 from 6.47. Probably this all has a formal presentation in terms of the direct sum and direct product subspace notations, but let's not do that now!

**Derive 3.11.** If we consider r,r' only from the set 1,2. then 3.11 says this

$$w^{\dagger r}(p)\, w^{r'}(p) = (E/mc^2)\, \delta_{r,r'}$$

For r = r' = 1, I have shown this to be true, and we use the fact that $p_z^2 + p_+p_- = p^2$. In fact, this fact is obviously true for all four columns dotted with themselves, so we know that

$$w^{\dagger r}(p)\, w^{r}(p) = (E/mc^2) \qquad r = 1,2,3,4$$

The following fact is also true

$$w^{\dagger r}(\varepsilon_r p)\, w^{r}(\varepsilon_r p) = (E/mc^2) \qquad r = 1,2,3,4$$

because for r = 3,4 we are negating p in both terms which makes no bilinear difference.

So the issue here is with the *orthogonality*! For this purpose, we can think of the vectors in the following simplified way:

$$w^1 = \begin{pmatrix} 1 \\ 0 \\ p_z \\ p_+ \end{pmatrix} \qquad w^2 = \begin{pmatrix} 0 \\ 1 \\ p_- \\ -p_z \end{pmatrix} \qquad w^3 = \begin{pmatrix} p_z \\ p_+ \\ 1 \\ 0 \end{pmatrix} \qquad w^4 = \begin{pmatrix} p_- \\ -p_z \\ 0 \\ 1 \end{pmatrix}$$

$$w^{1\dagger} = (1\ 0\ p_z\ p_-) \qquad w^{2\dagger} = (0\ 1\ p_+\ -p_z) \qquad w^{3\dagger} = (p_z\ p_-\ 1\ 0) \qquad w^{4\dagger} = (p_+\ -p_z\ 0\ 1)$$

By brute force I find these things to be true:

$w^{1\dagger}(p)\, w^{2}(p) = 0 + 0 + p_z\, p_- - p_z\, p_- = 0$
$w^{1\dagger}(p)\, w^{3}(-p) = p_z - p_z = 0$
$w^{1\dagger}(p)\, w^{4}(-p) = p_- - p_- = 0$
$w^{2\dagger}(p)\, w^{3}(-p) = p_+ - p_+ = 0$
$w^{2\dagger}(p)\, w^{4}(-p) = p_z - p_z = 0$
$w^{3\dagger}(p)\, w^{4}(p) = p_z\, p_- - p_z\, p_- = 0$ and of course $w^{3\dagger}(-p)\, w^{4}(-p) = 0$ as well.

So by brute force I have shown that everything is orthogonal as claimed in 3.11 and the diagonal values are also as claimed. So I have now proven 3.11 to be true by the brute force method.

Obviously there must be a simpler way to show this. Here was one attempt:

**********

Lemma: $\gamma^0\, w^r(p) = \varepsilon_r\, w^r(-p) \qquad r = 1,2,3,4) \qquad$ where $\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Proof: we know that $\gamma^0$ negates the signs of the lower two components of our spinor $w^r(p)$. For the first two columns in 3.7, we can undo this sign change by going $\mathbf{p} \rightarrow -\mathbf{p}$, so we can say

$$\gamma^0\, w^r(p) = w^r(-p) \qquad r = 1,2$$

For the last two columns, we can negate the first two components with $\mathbf{p} \rightarrow -\mathbf{p}$, then we can negate all four components with an overall minus sign, and together these two actions will compensate for the action of $\gamma^0$ on our spinor, so

$$\gamma^0 \; w^r(p) \; = \; -w^r(-p) \qquad r = 3,4$$

We can combine these by saying

$$\gamma^0 \; w^r(p) \; = \varepsilon_r \, w^r(-p) \qquad r = 1,2,3,4$$

Now consider

$$\varepsilon_r \, \delta_{r,r'} = w^{\dagger r}(p) \, \gamma^0 \; w^{r'}(p) \; = \varepsilon^r \; w^{\dagger r}(p) \, w^{r'}(-p)$$

which tells us that

$$w^{\dagger r}(p) \, w^{r'}(-p) \; = \delta_{r,r'}$$

This gives many of the desired results, but not all of them. This gives the desired answers for cross-group combinations, but *within* the group r ,r' = 1,2 or 3,4 this is not the desired result, though it is true. We have to show this

$$w^{\dagger r}(p) \, w^{r'}(+p) \; = \delta_{r,r'} \qquad\qquad r,r' \text{ in } 1,2 \text{ or in } 3,4$$

This is all true, but the only way I know to show it is brute force. Enough!

**********

**Comments $\psi_r(\mathbf{p})$:** Our full solution is

$$\psi^r(\mathbf{p}) \; = \exp(-i\varepsilon_r \, p_\mu x^\mu \, / \, \hbar \,) w^r(\mathbf{p})$$

What happens when we apply the 4-momentum *operator*?

$$\hat{p}_\mu \exp(-i\varepsilon_r \, p_\mu x^\mu/\hbar) = +i\hbar \, \partial^\mu \exp(-i\varepsilon_r \, p_\mu x^\mu/\hbar) \; = + \, \varepsilon_r p_\mu \exp(-i\varepsilon_r \, p_\mu x^\mu/\hbar \,)$$

Thus we have this slightly strange situation: (as if the neg energy solutions "go backwards")

$$\hat{p}_\mu \, \psi^r(\mathbf{p}) \; = \varepsilon_r p_\mu \, \psi^r(\mathbf{p})$$

If we look at the 3-momentum operator we have (both sides up now)

$$\hat{\mathbf{p}} \, \psi^r(\mathbf{p}) \; = \varepsilon_r \mathbf{p} \, \psi^r(\mathbf{p})$$

So the "negative energy" solutions not only have negative energy, they also have momentum $-\mathbf{p}$.

Parse sentence on page 31 middle of page:

(1) Equation 3.11 when applied to spinors of opposite energy says $w^{r\dagger}(\mathbf{p})w^{r'}(-\mathbf{p}) \stackrel{\bullet}{=} \delta_{r,r'}$. Assume that r' is the negative energy spinor. We know from above that $w^{r'}(-\mathbf{p})$ really has momentum $+\mathbf{p}$. So these two spinors here have opposite energy, but the both have momentum $+\mathbf{p}$ . Thus "two plane wave solutions of the same spatial momentum **p** but of opposite energy are orthogonal in the $w^{\dagger}w$ sense."

**Comments on page 32 equation A:** $\boldsymbol{\sigma}\bullet\mathbf{s}\ w = w$

First, think of this in a 2x2 sense. Here is what we would do:

$\boldsymbol{\sigma}\bullet\hat{\mathbf{z}}\ |\alpha\rangle = \sigma_3|\alpha\rangle = +1\ |\alpha\rangle$ where $|\alpha\rangle$ = column (1 0)

$\boldsymbol{\sigma}\bullet\hat{\mathbf{z}}\ |\beta\rangle = \sigma_3|\beta\rangle = -1\ |\beta\rangle$ where $|\beta\rangle$ = column (0 1)

Now start with the upper line and do this:

$$R\boldsymbol{\sigma}\bullet\hat{\mathbf{z}}\ R^{-1}R\ |\alpha\rangle = +1\ R\ |\alpha\rangle \quad \Rightarrow \quad \hat{\mathbf{z}}\bullet(R\boldsymbol{\sigma}R^{-1})\ |\theta,\varphi\rangle = +1\ |\theta,\varphi\rangle$$

where R= R(θ,φ) rotates the up state to wherever you want it. R = 2x2 matrix of SU(2). Then use our quoted fact above

$$R^{-1}\mathbf{V}\ R = R\ \mathbf{V}$$

and we have , where R is the usual 3x3 matrix of O(3),

$$\hat{\mathbf{z}}\bullet(R^{-1}\boldsymbol{\sigma})\ |\theta,\varphi\rangle = +1\ |\theta,\varphi\rangle$$

Now move to the other side to get

$$R\hat{\mathbf{z}}\bullet\boldsymbol{\sigma}\ |\theta,\varphi\rangle = +1\ |\theta,\varphi\rangle$$

And call this thing $\hat{\mathbf{s}} = R\hat{\mathbf{z}}$ and we get our desired result:

$\hat{\mathbf{s}}\bullet\boldsymbol{\sigma}\ |\theta,\varphi\rangle = +1\ |\theta,\varphi\rangle$ $|\theta,\varphi\rangle = R(\theta,\varphi)\ |\alpha\rangle$

which indicates a state where the spin is polarized in the $\hat{\mathbf{s}}$ direction. So here are the conclusions:

$\hat{\mathbf{s}}\bullet\boldsymbol{\sigma}\ |\theta,\varphi,\text{ up}\rangle = +1\ |\theta,\varphi,\text{ up}\rangle$ $|\theta,\varphi\rangle = R(\theta,\varphi)\ |\alpha\rangle$

$\hat{\mathbf{s}}\bullet\boldsymbol{\sigma}\ |\theta,\varphi,\text{ dn}\rangle = -1\ |\theta,\varphi,\text{ dn}\rangle$ $|\theta,\varphi\rangle = R(\theta,\varphi)\ |\beta\rangle$

We also know that $|\beta\rangle = R_y(\pi)\ |\alpha\rangle$ and etc from Levitt world. The point is that if you can make a state that is an eigenstate of $\hat{\mathbf{s}}\bullet\boldsymbol{\sigma}$, then that state has spin polarized in the $\hat{\mathbf{s}}$ direction (or opposite).

In the 4x4 world, we have

$$\mathbf{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$$

Based on this $\mathbf{\Sigma}$ and the above discussion, the spinors $w_{\mathbf{r}}(p)$ for r = 1,2,3,4 have spins +,–.+,– . However, we are going to define the v in such a way that when you list the spinors in their normal order, the spin sequence comes out being +,–.–,+. The goal is to change the spins of the negative energy solutions relative the *name* you use for the spinor. This happens soon below.

**The spin vector situation.** In a rest frame, we have $s^{\mu} = (0,\mathbf{s})$ and $p^{\mu} = (m,\mathbf{0})$. This suggests that we have in general $s_{\mu}\, p^{\mu} = 0$ as a world scalar, and also $s^{\mu}\, s_{\mu} = s^{j}s_{j} = -\, s^{j}s^{j} = -\mathbf{s}\bullet\mathbf{s} = -1$ , also a world scalar.

Now consider u(p,s) being a positive energy solution with rest spin **s**. Then we expect:

$\hat{\mathbf{s}} \bullet \mathbf{\Sigma}\, u(p,s) = u(p,s)$

$\hat{\mathbf{s}} \bullet \boldsymbol{\sigma}\, |\theta,\varphi, up\rangle = +1\ |\theta,\varphi, up\rangle$ // comparing to the above

How would you write out u(p,s) ? In rest frame we can do it this way

$$u(p,s) = R(\theta,\varphi)\, u(p,s = +z) = R(\theta,\varphi) \times R(\theta,\varphi) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

But the point is that we are defining u(p,s) to be positive energy and s = **s** in rest frame, but of course things are 4-vectors. So if $\hat{z} = (0,0,1)$ we have $u_{\mathbf{z}}^{\mu} = (0,0,0,1)$ . Then my conjecture above says

$$u(p,u_{\mathbf{z}}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad u(p,-u_{\mathbf{z}}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = R_{\mathbf{y}}(\pi) \times R_{\mathbf{y}}(\pi) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Here I am seeing a suggestion of the **(1/2,0) ⊕ (0,1/2)** sense of this Dirac representation. Of course I have written the above in the rest frame. You can then boost out so that $u(p,u_{\mathbf{z}}) = w^{\mathbf{1}}(p)$ to any p you want. You can go to the rest frame to observe where **s** is pointing. Of course in a boosted frame, it gets distorted into some $s^{\mu}$.

### 3.2 Projection Operators for Energy and Spin

**Energy Projectors**.

These are easy and obvious. We have

$\Lambda_{\pm} = (\pm\not{p} + mc)/(2mc)$.

This works because we know that $\not{p} - mc = 0$ for r=1 and 2, meaning $\not{p} = mc$ when applied to 1 or 2. Thus,( $\not{p} + mc)/(2mc) = ( mc + mc)/(2mc) = 1$ as an operator when going on + states. And for negative energy states, we know that $\not{p} = - mc$ so again we get 1 when acting on them with $\Lambda_-$.

**Spin projectors.**

**Comments on gamma matrices.** The four basic $\gamma^\mu$ are the originally defined ones (in terms of $\alpha_i$ and $\beta$). We are then allowed to *define* lower index ones by $\gamma_\mu = g_{\mu\nu}\gamma^\nu$ even though we would not say $\gamma^\mu$ was a 4-vector by itself. But we know that $\gamma^\mu$ *does* transform as a 4-vector when sandwiched between $\bar{\psi}$ and $\psi$. So this extension allows us to say things like $j^\mu = \bar{\psi}\,\gamma^\mu\,\psi \;\Rightarrow j_\mu = \bar{\psi}\,\gamma_\mu\,\psi$ . [ I later learned that $\gamma^\mu$ transforms as a 4-vector on the action of Dirac rep xforms, just as $p^\mu$ does on Vector rep xforms, see LG doc.]

The various *other* gamma matrices do not vary up or down and probably you should always write them the same way. Looking on page 25, we see that you might be tempted to have $\gamma_5 = -\gamma^5$ from the way $\gamma_5$ is defined, but B&D go out of their way to clearly state that $\gamma_5 = \gamma^5$ by definition! Later we encounter a $\gamma^6$ which is the same idea $\gamma_6 = \gamma^6$.

So the moral here is to be careful with $\gamma^3 = -\gamma_3$ in the spin discussion below. For example, when we have the notation $\not{u}_z$ the valuation in the rest frame would be $+\gamma_3 s^3 = +\gamma_3$ since $u_z^\mu = (0,0,0,1)$

Consider the combination $\gamma_5\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}$ so that $\gamma_5\gamma_3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix}$. What does $\gamma_5\gamma_3$ do to our four basic spinors?

$$\begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix} w^1 = +\, w^1$$
$$\begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix} w^2 = -\, w^2$$
$$\begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix} w^3 = -\, w^3$$
$$\begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix} w^4 = +\, w^4$$

These four signs *agree* with 3.16 and are bringing out the "sign of the spin" in the four cases. Consider then (ignore right two columns below for the moment)

| | | | |
|---|---|---|---|
| $\frac{1}{2}(1 + \gamma_5\gamma_3)\; w^1 = w^1$ | spin up | $\Sigma(u_z)\, w^1 = w^1$ | $\Sigma(u_z)\, u(u_z) = u(u_z)$ |
| $\frac{1}{2}(1 - \gamma_5\gamma_3)\; w^2 = w^2$ | spin down | $\Sigma(-u_z)\, w^2 = w^2$ | $\Sigma(-u_z)\, u(u_z) = 0$ |
| $\frac{1}{2}(1 - \gamma_5\gamma_3)\; w^3 = w^3$ | spin down | $\Sigma(-u_z)\, w^3 = w^3$ | $\Sigma(u_z)\, v(u_z) = v(u_z)$ |
| $\frac{1}{2}(1 + \gamma_5\gamma_3)\; w^4 = w^4$ | spin up | $\Sigma(u_z)\, w^4 = w^4$ | $\Sigma(-u_z)\, v(u_z) = 0$ |

Here we see the sign sequence noted above for the $w_r$ spinors being +, –, –, +. Now, when we define the u and v, we define $w^3(p) = v(p, -u_z)$ and then when you list the states in the following order, you get the following spin signs:

u(p,s) +
u(p,-s) –
v(p,s) +
v(p,-s) –

These are the rest frame projections we want. So write

$$\Sigma(u_z) = \tfrac{1}{2}(1 + \gamma_5\gamma_3) = \tfrac{1}{2}(1 + \gamma_5 \not{u}_z)$$

and we generalize to

$$\Sigma(s^\mu) = \tfrac{1}{2}(1 + \gamma_5 \not{s}) \qquad \text{for an arbitrary spin 4-vector.}$$

We then get, in the rest frame, the third column above. But in terms of the u and v states, we get the far right column, and we see how the minus of the definition of v cancels out the minus of $\Sigma(-u_z)$ $w^3 = w^3$. This rightmost column is equations D on page 34.

The point of 3.22 is that the spin-projectors are in "covariant form" just the way $(\not{p}-m)\psi = 0$ is in covariant form. That is to say, we had earlier that

$$S(a)\,\not{\partial}\,S^{-1}(a) = S(a)\,\gamma^\mu\,S^{-1}(a)\,\partial_\mu = S(a)\,\gamma^\mu\,S^{-1}(a)\,(a_\nu{}^\mu \partial'^\nu) = \{\,S(a)\,\gamma^\mu\,S^{-1}(a)\,a_\nu{}^\mu\,\}\,\partial'^\nu = \not{\partial}'$$

The slashed object is taken into the new reference frame by the transformation S. What happens here when we do this with a non-operator 4-vector? Same thing. That is to say, for example,

$$S(a)\,\not{p}\,S^{-1}(a) = \not{p}' \qquad \text{where p is just a 4-vector}$$

So the same thing will happen with ANY operator when contains things like $\not{p}$ or $\not{s}$. The equation in question will have the same form in primed as in unprimed, where you are acting on some $\psi(x)$ so that you have $\psi' = S(a)\,\psi$. So that is what they mean when they say $\frac{1}{2}(1 + \gamma^5\not{s})$ is in covariant form. They don't mean it is a world scalar. So this is why we always try to find a covariant form in the rest frame, then we can take it to any frame in this way. In other words, going from page 34 D to 3.22.

Comment: I think this all is nothing more than a naming of states. As shown in 3.15, we define v(p,s) to have spin down.

**Clarifications added 10.11.08**

(1) Question: Is $\not{p}$ a "world scalar" in some sense, or is it not?

In pondering this question, I roll out lots of our now known technology, prove covariance of the Dirac equation two ways, and then state a solid conclusion at the end!

I now know the following facts:

$$S^{-1}\,\gamma^\alpha\,S = \Lambda^\alpha{}_\nu\,\gamma^\nu \qquad \gamma^\mu \text{ transforms as a 4-vector object under Dirac rep action}$$

$\Lambda^{-1}\, p^\mu \Lambda = \Lambda^\mu{}_\nu p^\nu$ $\quad$ $p^\mu$ transforms as a 4-vector object under Vector rep action

Here U and $\gamma^\alpha$ are "operators" in the Dirac 4x4 space, whereas $\Lambda$ and p are operators in the $\infty$x$\infty$ QM HS. The following two facts are true

$$S^{-1}\, \not{p}'\, S = p'_\mu\, S^{-1}\, \gamma^\mu\, S = p'_\mu\, \Lambda^\mu{}_\nu\, \gamma^\nu = p_\nu\, \gamma^\nu = \not{p} \qquad p' = \Lambda p$$

$$\Lambda^{-1}\, \not{p}\, \Lambda = \gamma_\mu\, \Lambda^{-1}\, p^\mu\, \Lambda = \gamma_\mu\, \Lambda^\mu{}_\nu p^\nu = \gamma_\mu\, p'_\mu = \not{p}'$$

where we use these facts from our curvilinear doc

$$x'^a = T^a{}_b\, x^b \qquad x'_a = T_a{}^b\, x_b$$
$$x^a = T_b{}^a\, x'^b = x'^b\, T_b{}^a \qquad x_a = T^b{}_a\, x'_b = x'_b\, T^b{}_a$$

That is to say, we know these two facts

$$S^{-1}\, \not{p}'\, S = \not{p} \qquad p' = \Lambda p$$
$$\Lambda^{-1}\, \not{p}\, \Lambda = \not{p}'$$

A. Dirac covariance in coordinate space

Now let's look very carefully at the covariance of the Dirac equation:

$$(\gamma^\mu \hat{p}_\mu - m)\psi(x) = 0 \qquad \text{frame F}$$
$$(\gamma^\mu \hat{p}'_\mu - m)\psi'(x') = 0 \qquad \text{frame F'}$$

Start with $\psi'(x') = S\, \psi(x)$ and start with the second equation above

1 $\quad (\gamma^\mu \hat{p}'_\mu - m)\psi'(x') = 0$

2 $\quad (\gamma^\mu \hat{p}'_\mu - m)\, S\, \psi(x) = 0$

Here we have to think of $x = \Lambda^{-1}x'$ inside $\psi(x)$, so $\hat{p}'_\mu$ does not just give zero when acting on $S\, \psi(x)$. In fact, we know that

$$\partial_\mu{}' f(x) = \frac{\partial f}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\, \frac{\partial f}{\partial x^\nu} = \Lambda_\mu{}^\nu\, \partial_\nu f(x) \qquad \Rightarrow \qquad \hat{p}'_\mu\, f(x) = \Lambda_\mu{}^\nu\, \hat{p}_\nu\, f(x)$$

which we can *abbreviate* as $\hat{p}'_\mu = \Lambda_\mu{}^\nu\, \hat{p}_\mu$, just as if p were a 4-vector non-operator object. Of course this really *means* the equality is true *when* we act of a function of x where $x = \Lambda^{-1}x'$ . Applying this idea to the linear combination $S\, \psi(x)$ we arrive at

3 $\quad (\gamma^\mu\, \Lambda_\mu{}^\nu\, \hat{p}_\nu - m)\, S\, \psi(x) = 0$

Now apply $S^{-1}$ from the left and shuffle the S around a bit to get

4 $S^{-1} (\gamma^\mu \Lambda_\mu{}^\nu \hat{p}_\mu - m) S \psi(x) = 0$

5 $( [S^{-1}\gamma^\mu S] \Lambda_\mu{}^\nu \hat{p}_\mu - m) \psi(x) = 0$

Now use $S^{-1} \gamma^\mu S = \Lambda^\mu{}_\alpha \gamma^\alpha$ which we stated above to get

6 $(\Lambda^\mu{}_\alpha \gamma^\alpha \Lambda_\mu{}^\nu \hat{p}_\mu - m) \psi(x) = 0$

Then use the fact that $\Lambda^\mu{}_\alpha \Lambda_\mu{}^\nu = \delta_\alpha{}^\nu$ for any $T^a{}_b$ type transformation and we get

7 $(\gamma^\mu \hat{p}_\mu - m) \psi(x) = 0$

Thus we have shown that 1 => 7 and the Dirac equation is "covariant".

<u>B. Dirac covariance in coordinate space</u>

Now let's repeat all of this using bra-ket notation. Start in the same place as above

1 $(\gamma^\mu \hat{p}'_\mu - m)\psi'(x') = 0$ or $(\gamma^\mu \hat{p}'_\mu - m)\langle x'|\psi'\rangle$ or $(\hat{p}'_\mu - m)\langle x'|\psi'\rangle$

Replace $\hat{p}'_\mu \langle x'| f\rangle = \langle x'|P^\mu| f\rangle$ in the usual manner to get ( we used $P = P^\dagger$ )

a $\langle x'| (\gamma_\mu P^\mu - m)|\psi'\rangle = 0$ or $\langle x'| (\not{P} - m)|\psi'\rangle = 0$

Now expose $\Lambda^\dagger$ on the left side, and insert a pair prior to the ket to get ( $\Lambda$ in the QM HS is unitary)

b $\langle x| \Lambda^\dagger (\gamma_\mu P^\mu - m) \Lambda \Lambda^\dagger |\psi'\rangle = 0$

Now use "explanation of...." Section 4.1 which says $R^\dagger|\psi'_r\rangle = S(R)_{rs} | \psi_s(p)\rangle$ to get

c $\langle x| \Lambda^{-1} (\gamma_\mu P^\mu - m) \Lambda S |\psi\rangle = 0$

d $\langle x| (\gamma_\mu [\Lambda^{-1} P^\mu \Lambda] - m) S |\psi\rangle = 0$

Next, use our rule $\Lambda^{-1} p^\mu \Lambda = \Lambda^\mu{}_\nu p^\nu$ given above for transforming a vector operatort in QM HS

e $\langle x| (\gamma_\mu \Lambda^\mu{}_\nu P^\nu - m) S |\psi\rangle = 0$

Now apply $S^{-1}$ to the left of both sides

f $S^{-1}\langle x| (\gamma_\mu \Lambda^\mu{}_\nu P^\nu - m) S |\psi\rangle = 0$

g $\langle x| ([S^{-1} \gamma_\mu S] \Lambda^\mu{}_\nu P^\nu - m) |\psi\rangle = 0$

Now apply our rule that $S^{-1} \gamma^\mu S = \Lambda^\mu{}_\alpha \gamma^\alpha$ which says $S^{-1} \gamma_\mu S = \Lambda_\mu{}^\alpha \gamma_\alpha$

h $\langle x| (\Lambda_\mu{}^\alpha \gamma_\alpha \Lambda^\mu{}_\nu P^\nu - m) |\psi\rangle = 0$

Then use the orthogonality as usual to get

i or $\langle x| (\not{P} - m) |\psi\rangle = 0$

now replace $\langle x| P^\nu |f\rangle$ by $\hat{p}^\nu \langle x|f\rangle$ to et

7 $(\gamma_\nu \hat{p}^\nu - m) \langle x| \psi\rangle = 0$ or $(\hat{\not{p}} - m) \psi(x) = 0$

So here we have two separate "proofs" of the covariance of the Dirac equation. The first is done with wavefunctions, the second is done in the QM HS. In the second proof, both the $\Lambda$ and S transformations are visible. That is, we encounter both $[\Lambda^{-1} P^\mu \Lambda]$ in the QM HS, and we encounter $[S^{-1} \gamma_\mu S]$ in the Dirac space, and we have to use both these transformation rules. We can now look back at some of the stages of this second proof:

a $\langle x'| (\gamma_\mu P^\mu - m)|\psi'\rangle = 0$ =>

c' $\langle x| S^{-1}\Lambda^{-1} (\gamma_\mu P^\mu - m) \Lambda S |\psi\rangle = 0$ =>

i' $\langle x| \gamma_\mu P^\mu - m) |\psi\rangle = 0$

so if we look just at the operator center core here, we see this fact

$S^{-1}\Lambda^{-1} (\gamma_\mu P^\mu) \Lambda S = \gamma_\mu P^\mu$ or $U^{-1}(\gamma_\mu P^\mu) U = \gamma_\mu P^\mu$

or

$U^{-1}\not{P}U = \not{P}$ where $U = \Lambda S$

So this now brings us to the answer to an implied question " is $\not{p}$ a world scalar? " Here we see that under the *combined* Lorentz tranformations $\Lambda$ (which acts on QM HS operators and so acts on p) and S (which acts in the Dirac spinor space and so acts on $\gamma$), the operator object $\not{P}$ is indeed a scalar. Under either of these transformations alone, we find something that is not a scalar.

Now if we have some non-operator 4-vector $p^\mu$, then in one Lorentz frame we could compute $\not{p}$ and we would get a certain 4x4 matrix. If we go to another Lorentz frame and compute $\not{p}'$, we will get a *different* 4x4 matrix if we use the same $\gamma$ matrices, just treating them as constants. So in that sense, one would never say $\not{p}$ was a world-scalar. This is the usual meaning of $\not{p}$ wrt LT's. That is, we transform the p but not the $\gamma$.

Consider the following which shows that s is in fact a "world scalar matrix" :

$$s' = p'_\mu \gamma'^\mu = (\Lambda^{-1} p_\mu \Lambda)\,(S^{-1} \gamma^\mu S) = \Lambda_\mu{}^\nu p_\nu \; \Lambda^\mu{}_\alpha \gamma^\alpha = p_\mu \gamma^\mu = s$$

which we could write as

$$(\not{p})' = \not{p} \qquad \text{where } (\not{p})' \equiv p'_\mu \gamma'^\mu$$

But when we write $\not{p}'$ , we always mean $p'_\mu \gamma^\mu$ so we then have

$$\not{p}' \neq \not{p} \qquad \text{and in fact} \qquad \not{p}' = p'_\mu \gamma^\mu = \Lambda_\mu{}^\nu p_\nu \gamma^\mu \neq \not{p}$$

So without any confusion, we would say that what we always denote by $\not{p}$ is not a world scalar.

(2) How is the spin operator defined and why?

This note concerns spin and the spin projection operator. I think one should realize that the association of "spin up" with $w^4$ is an arbitrary decision made by B&D, it is not forced by anything. Exactly why they do this, I comment at the end of this note, but the fact is that they arbitrarily do it.

They *could have* done it differently. They could have let $\mathbf{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$ be the "spin operator" (all indices are down here, $\Sigma_i$ and $\sigma_i$) . It is possible to write $\mathbf{\Sigma} = \gamma_6 \boldsymbol{\gamma}$ so that $\Sigma_3 = \gamma_6 \gamma^3$. Our spin projector for aligned spin-up states would have been

$$\Lambda^+ = \tfrac{1}{2}(1 + \mathbf{\Sigma}_3) = \tfrac{1}{2}(1 + \gamma_6 \boldsymbol{\gamma}^3)$$

We could then write, in this rest frame, that $\not{s} = s^\mu \gamma_\mu = s^3 \gamma_3 = \gamma_3 = -\gamma^3$ to get

$$\Lambda^+ = \tfrac{1}{2}(1 + \mathbf{\Sigma}_3) = \tfrac{1}{2}(1 + \gamma_6 \boldsymbol{\gamma}^3) = \tfrac{1}{2}(1 - \gamma_6 \not{s})$$

and then the covariant projector for any spin would have been

$$\Lambda^+ = \tfrac{1}{2}(1 - \gamma_6 \not{s})$$

<u>But this is not what they did</u>. Instead, they use $\gamma_5 \gamma_3 = -\gamma_5 \gamma^3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix}$ or $\mathbf{\Sigma}' \equiv -\gamma_5 \boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix}$ as the "spin operator". The difference between this and the above is the minus sign on the lower $\boldsymbol{\sigma}$. In the rest frame, the spin up projector would be

$$\Lambda^+ = \tfrac{1}{2}(1 + \mathbf{\Sigma}'_3) = \tfrac{1}{2}(1 - \gamma_5 \gamma^3)$$

By the same argument as above, this then leads to the following covariant projector

$$\Lambda^+ = \tfrac{1}{2}(1 + \gamma_5 \not{s})$$

Again, I think this definition of the "spin operator" and its resultant covariant form is arbitrary, but they don't ever say that in the book.

One motivation for doing what they did is this: if we want to associate a physical positron going forward in time with a negative energy electron going backward in time, then going backward in time ought to flip the spin direction if we think of spin as an "angular momentum". If you have a rotating object, time reversal makes it rotate backwards, and the angular mometum vector reverses. So this might be the reason they want to associate v(p,s) with $w^4$ instead of $w^3$. We know that $w^4$ has spin down and we know it is a negative energy electron state, so the physical positron associated with it should have spin up, and we cause this to happen by the way we define the spin operator. That is to say, we think of v(p,s) = $w^4$ as a spin up positron state, meaning it is spin up when we are in the rest frame and set $\hat{s} = \hat{z}$ . In this same manner, we think of u(p,s) as a spin up physical electron.

**Combination Projectors**

These are stated on page 35. If you have some linear combination of the $w^r$, you can project out what you want. For example:

$$f = a_1 w^1 + a_2 w^2 + a_3 w^3 + a_4 w^4$$

$$P_3 f = a_3$$

Recall that state $w^3$ has negative energy and negative spin. Notice he has $\Sigma(u_z)$ where

$u_z{}^\mu = a^\mu{}_\nu \, u_{0z}{}^\nu$ where $u_{0z}{}^\nu = (0,0,0,1)$ which is the $\hat{z}$ thing.

So equations page 35 A *are* in some general frame, not just the rest frame.

**Prove that $[ \Sigma(s), \Lambda_\pm ] = 0$ when s.p = 0.**

This is easy. Install all the pieces, and we have to show that

$[\gamma_5 \not{s}, \not{p}] = 0$ or $[ \gamma_5 \gamma^\mu, \gamma^\nu] s_\mu p_\nu = 0$

But $[ \gamma_5 \gamma^\mu, \gamma^\nu] = \gamma_5 \{ \gamma^\mu, \gamma^\nu\} = \gamma_5 \, g^{\mu\nu}$ so we get

$[ \gamma_5 \gamma^\mu, \gamma^\nu] s_\mu p_\nu = 0 \iff \gamma_5 \, g^{\mu\nu} s_\mu p_\nu = 0$ or s.p = 0, as claimed.

**3.3 Physical Interpretation of Free-particle Solutions and Packets (35)**

The expansion 3.23 includes the r = 1,2 positive energy states only, hence the super (+) on the ψ. The phase space factors do indeed result in 3.24 which then tells you that b(p,s) is the normalized momentum space wavefunction which you square to get probability. Note that b is just a coefficient, possibly complex, it is not a creation or destruction operator! [ but I am sure it will becomes such eventually ] .

The big next step is going to be computation of the 4-current for this packet!

First, what about the zeroth component? The current we know is

$j^\mu = c\,\bar\psi\,\gamma^\mu\,\psi$ so $j^0 = c\,\bar\psi\,\gamma^0\,\psi = c\,\psi^\dagger\psi$ $J^\mu = \int d^3x\, j^\mu$

The total 0th component of the current is then $c\int d^3x\,\psi^\dagger\psi$ so we have

$$J^0 = c\int d^3x\,\psi^\dagger\psi = c * (3.24)$$

so this $J^0$ is what we have already computed (and we are reminded that it is positive definite). But now, let's do the entire 4-vector using the **Gordon (1928) Decomposition** 3.26 to get something like 3.28. So now we are doing "**derive 3.28**" but with μ. [ the decomposition is derived at the top of page 37 for us. ]

$$J^\mu = c\int d^3x\;\bar\psi\,\gamma^\mu\,\psi$$

Now, let's use 3.23 as is for ψ, and then prime all the integration/sum variables for $\bar\psi$ and we get this, where I will attempt some shorthand: ($f = 1/(2\pi\hbar)^3\, mc^2/\sqrt{EE'}$ ) ,

$$J^\mu = c\int d^3x \int d^3p'\int d^3p\, f\,\Sigma_{s,s'} b^{*\prime}\, u^{\dagger\prime}\, e^{+ip'\cdot x}\,\gamma^0\gamma^\mu\, b\, u\, e^{-ip\cdot x}$$

$$= c\int d^3x \int d^3p'\int d^3p\, f\,\Sigma_{s,s'} b^{*\prime}\, b\,\bar u'\,(\, e^{+ip'\cdot x}\,\gamma^\mu\, e^{-ip\cdot x})\, u$$

$$= \int d^3x \int d^3p'\int d^3p\, f\,\Sigma_{s,s'} b^{*\prime}\, b\,\bar u'\,(c\,\bar\psi_2(x)\gamma^\mu\,\psi_1(x))\, u$$

where now the object $(\, e^{+ip'\cdot x}\,\gamma^0\gamma^\mu\, e^{-ip\cdot x})$ can subject itself to the Gordon decomposition (and here I have temporarily redefined the meaning of ψ just to fit the Gordon formula)

$\psi_1(x) = e^{-ip\cdot x/\hbar}\, u(p,s)$

$\psi_2(x) = e^{-ip'\cdot x/\hbar}\, u(p',s')$ $\psi_2(x)^\dagger = u'^\dagger\, e^{+ip'\cdot x/\hbar}$ $\bar\psi_2(x) = \bar u'\, e^{+ip'\cdot x/\hbar}$

So now we want to apply 3.26 to:

$$c\bar\psi_2(x)\,\gamma^\mu\,\psi_1(x)$$

We will need these computations:

$\hat p^\mu\,\psi_1(x) = i\hbar\partial^\mu\,(e^{-ip\cdot x/\hbar})u = +p^\mu\,(e^{-ip\cdot x/\hbar})u = +p^\mu\,\psi_1(x)$

$\hat p^\mu\,\bar\psi_2(x) = i\hbar\partial^\mu\,(e^{+ip'\cdot x/\hbar})\bar u' = -\,p'^\mu\,(e^{-ip\cdot x/\hbar})\,\bar u' = -\,p'^\mu\bar\psi_2(x)$

$\hat p_\nu(\,\bar\psi_2(x)\,\sigma^{\mu\nu}\,\psi_1(x)) = (\hat p_\nu(\,\bar\psi_2(x))\,\sigma^{\mu\nu}\,\psi_1(x) + \bar\psi_2(x)\,\sigma^{\mu\nu}\,(\hat p_\nu\,\psi_1(x)\,)$

$= -\,p'_\nu\bar\psi_2(x)\,\sigma^{\mu\nu}\,\psi_1(x) + \bar\psi_2(x)\,\sigma^{\mu\nu}\,p_\nu\,\psi_1(x)$

$$=(-\,p'_\nu + p_\nu)\,\bar\psi_2(x)\,\sigma^{\mu\nu}\,\psi_1(x)$$

So we get

$$c\bar\psi_2(x)\,\gamma^\mu\,\psi_1(x) = (1/2m)\,\{\ \bar\psi_2(x)\,p^\mu\,\psi_1(x) + p'^\mu\,\bar\psi_2(x)\,\psi_1(x) + i\ (+\,p'_\nu - p_\nu)\,\bar\psi_2(x)\,\sigma^{\mu\nu}\,\psi_1(x)\ \}$$

Now replace ψ's on the right. We can combine the two exponentials into $e^{i(p'-p)\cdot x/\hbar}$ and we have

$$c\bar\psi_2(x)\,\gamma^\mu\,\psi_1(x) = (1/2m)\,e^{i(p'-p)\cdot x/\hbar}\ \ \bar u'\,\{\ p^\mu + p'^\mu + i\ (+\,p'_\nu - p_\nu)\,\sigma^{\mu\nu}\ \}\,u$$

$$= (1/2m)\,e^{i(p'-p)\cdot x/\hbar}\ \ \bar u'\{\ p^\mu + p'^\mu + i\ (+\,p'_\nu - p_\nu)\,\sigma^{\mu\nu}\ \}\,u$$

Inserting this into the result above gives,

$$= \int d^3x \int d^3p' \int d^3p\ f\,\Sigma_{s,s'}\,b^{*\prime}\,b\ \ (c\,\bar\psi_2(x)\gamma^\mu\,\psi_1(x))$$

$$= \int d^3x \int d^3p' \int d^3p\ f\,\Sigma_{s,s'}\,b^{*\prime}\,b\,(1/2m)\,e^{i(p'-p)\cdot x/\hbar}\ \bar u'\,\{\ p^\mu + p'^\mu + i\ (+\,p'_\nu - p_\nu)\,\sigma^{\mu\nu}\ \}\,u$$

$$= \int d^3x \int d^3p' \int d^3p\ (f/2m)\,e^{i(p'-p)\cdot x/\hbar}$$
$$\Sigma_{s,s'}\,b^{*\prime}\,b\,\bar u'\,\{\ p^\mu + p'^\mu + i\ (+\,p'_\nu - p_\nu)\,\sigma^{\mu\nu}\ \}\,u$$

where $f = 1/(2\pi\hbar)^3\ mc^2/\sqrt{EE'}$ from above. **This agrees with the first long line in 3.28.**

Now we do the $d^3x$ integral to pick up $(2\pi\hbar)^3\,\delta^3(\mathbf{p}'-\mathbf{p})$ and E = E' we are then left with

$$\int d^3p\ (F/2m)\ \ \Sigma_{s,s'}\,b^{*\prime}\,b\,\bar u'\,\{\ p^\mu + p'^\mu + i\ (+\,p'_\nu - p_\nu)\,\sigma^{\mu\nu}\ \}\,u$$

where now p = p' and $F = mc^2/E$ and then $(F/2m) = c^2/(2E)$ so result is

$$c^2 \int d^3p(1/2E)\ \ \Sigma_{s,s'}\,b^*(p,s')\,b(p,s)\,\bar u\,(p,s')\,\{\ p^\mu + p^\mu + i\ (+\,p_\nu - p_\nu)\,\sigma^{\mu\nu}\ \}\,u(p,s)$$

It appears that the $\sigma^{\mu\nu}$ term goes away, and the first two terms add so we have

$$c^2 \int d^3p(1/E)\ \ \Sigma_{\pm s}\ b(p,s)\,p^\mu\,\Sigma_{\pm s'}\ b^*(p,s')\ \bar u\,(p,s')\,u(p,s)$$

Now let's recast (3.9b) into the u-spinor language [ this is row vector contracted against column ]

$$\bar w^r(p)\,w^{r'}(p) = \varepsilon_r\,\delta_{r,r'} \qquad\qquad r,r' = 1,2,3,4$$

which of course implies that

$$\bar{w}^{r}(p)\, w^{r'}(p) = \delta_{r,r'} \qquad r,r' = 1,2$$

which says that ( where f is some arbitrary function)

$$\Sigma_{r=1,2}\ f(r')\, \bar{w}^{r}(p)\, w^{r'}(p) = \Sigma_{r=1,2}\, f(r')\, \delta_{r,r'} = f(r) \qquad \text{for either value of } r'$$

which we can translate into

$$\Sigma_{\pm s}\ f(s')\, \bar{u}(p,s')\, u(p,s) = f(s) \qquad \text{for either value of } s$$

Now we think of f(s') = b*(p,s') and our result above becomes,

$$c^2 \int d^3p\,(1/E)\ \Sigma_{\pm s}\ b(p,s)\, p^{\mu}\, \Sigma_{\pm s'}\ b^*(p,s')\ \bar{u}\,(p,s')\, u(p,s)$$

$$= \int d^3p\,(p^{\mu} c^2/E)\ \Sigma_{\pm s}\ |b(p,s)|^2$$

And here is our full result then, where $\psi$ =

$$J^{\mu} = c \int d^3x\ \bar{\psi}^{(+)}\, \gamma^{\mu}\, \psi^{(+)} = \int d^3p\,(p^{\mu} c^2/E)\ \Sigma_{\pm s}\ |b(p,s)|^2$$

and from this we can set $p^0 = E/c$ to get

$$J^0 = c \int d^3x\ \bar{\psi}^{(+)}\, \gamma^0\, \psi^{(+)} = c \int d^3p\ \Sigma_{\pm s}\ |b(p,s)|^2$$

$$J^i = c \int d^3x\ \bar{\psi}^{(+)}\, \gamma^i\, \psi^{(+)} = \int d^3p\,(p^i c^2/E)\ \Sigma_{\pm s}\ |b(p,s)|^2$$

$$\mathbf{J} = c \int d^3x\ \bar{\psi}^{(+)}\, \boldsymbol{\gamma}\, \psi^{(+)} = \int d^3p\,(\mathbf{p}\, c^2/E)\ \Sigma_{\pm s}\ |b(p,s)|^2$$

where notice that the vector $\boldsymbol{\gamma}$ refers to $\gamma^i$. Now we can write the above as

$$J^0 = \langle \psi^{\dagger(+)} |\ c\ |\psi^{(+)}\rangle = c \int d^3p\ \Sigma_{\pm s}\ |b(p,s)|^2 = c \qquad \text{// normed in 3.24}$$

$$\mathbf{J} = \langle \psi^{\dagger(+)} |\ c\ \gamma^0\ \boldsymbol{\gamma}\ |\psi^{(+)}\rangle = \int d^3p\,(\mathbf{p}\, c^2/E)\ \Sigma_{\pm s}\ |b(p,s)|^2 = \langle c\ \gamma^0\ \boldsymbol{\gamma}\rangle = \langle c\ \boldsymbol{\alpha}\rangle$$

**and now we have derived the second line in 3.28.**

In coordinate space, we imagine $1 = \int d^3x \, \Sigma_{\pm s} \, | x,s><x,s|$ as our completeness. In momentum space we have to use $1 = \int d^3p \, \Sigma_{\pm s} \, | p,s><p,s|$ I guess (no factor). In momentum space, each u(p,s) component has amplitude b(p,s) and probability $|b(p,s)|^2$ . So yes, we have

$$\mathbf{J} = < \psi^{\dagger (+)} \, | \, c \, \gamma^0 \, \boldsymbol{\gamma} \, |\psi^{(+)}>_{\mathbf{x}} \; = < \mathbf{p} \, c^2/E>_{\mathbf{p}}$$

Recall from relativisitics that $p^0 = E/c$ and $\mathbf{p} \; = \gamma m \, d\mathbf{r}/dt \; = \gamma m \mathbf{v} = \gamma \boldsymbol{\beta} mc$ and $E = \gamma mc^2$ so we have

$$\mathbf{p} \, c^2/E \; = \gamma m \mathbf{v} \, c^2 \, / \gamma mc^2 = \mathbf{v}$$

So our big result is this:

$$\mathbf{J} = < \mathbf{v} >_{\mathbf{p}} \quad = \quad < \psi | \, \mathbf{v} \, | \psi > / < \psi | \psi > = < \psi | \, \mathbf{v} \, | \psi >$$

I guess when you have a packet, this thing here is called the **group velocity** of the packet, and you have to compute it and we see that it is just this:

$$\int d^3p (\mathbf{p} \, c^2/E) \; \Sigma_{\pm s} \; |b(p,s)|^2$$

Whatever it is, that is how fast the packet moves, ie, that is how fast your electron is moving.

Now, recall that

$j^{\mu} (x) = c \, \bar{\psi}(x) \, \gamma^{\mu} \, \psi(x)$ = " the 4-current density"

$J^{\mu} = c \int d^3x \; \bar{\psi}^{(+)} \, \gamma^{\mu} \, \psi^{(+)}$ = the 4-current" of the positive-energy packet"

$$= \int d^3x \, j^{\mu} (x) \quad = \quad \int d^3p (p^{\mu} \, c^2/E) \; \Sigma_{\pm s} \; |b(p,s)|^2 \quad = \quad \int d^3p \, v^{\mu} \; \Sigma_{\pm s} \; |b(p,s)|^2$$

So I would not refer to $J^{\mu}$ as "the average current for an arbitrary packet...". I would say that "the four current $J^{\mu}$ of the packet" was a weighted average of the velocity of each momentum-space component of the packet", and this weighted average is known as "the group velocity" of the packet. So in the sentence below 3.29 I think they have something jumbled.

B&D now point out this distinction between rel and non-rel theories:

(1) non-rel theory, $[\mathbf{p},H] = 0$ and $\mathbf{v} = \mathbf{p}/m$ is a constant of the motion.

(2) rel theory, $[\boldsymbol{\alpha},H] \neq 0$ so the velocity operator $\mathbf{v}_{\mathbf{op}} = c\boldsymbol{\alpha}$ is not a constant of the motion.

I am missing the point made here. How would you make an eigenfunction of c$\boldsymbol{\alpha}$? Ah, OK, that is the point. You must linearly combine multiple eigenstates of H to get eigenstate of cα because $[\boldsymbol{\alpha},H] \neq 0$. If this were 0, you could diagonalize both at the same time. So "linearly combine" is no doubt going to mean that you have to mix states of both positive and negative energies to get an eigenstate of c$\boldsymbol{\alpha}$.

**Page 38 we forge ahead:**

Now, in 3.30 we **add in some v(p,s) negative energy spinor action** with coefficient d*(p,s) and with $f = 1/(2\pi\hbar)^3\, mc^2/\sqrt{EE'}$ .

$$J^k = c\int d^3x \int d^3p' \int d^3p\, f\, \Sigma_{s,s'}$$

$$\{\, b^{*\prime}\, \bar{u}'\, e^{+ip'\cdot x} + d'\, \bar{v}'\, e^{-ip'\cdot x} \,\}\, \gamma^k \,\{ b\, u\, e^{-ip\cdot x} + d^*\, v\, e^{+ip\cdot x} \,\}$$

Now we have 4 terms instead of just 1 term. OK, it would take me about 8 pages to derive 3.31, so instead of doing that, let's take note of the main points which are these:

1) in the first cross term $b^{*\prime}\, \bar{u}'\, e^{+ip'\cdot x}\, \gamma^k\, d^*\, v\, e^{+ip\cdot x}$ we have the same sign on the expo factors. The Delta function $\delta(\mathbf{p}+\mathbf{p}')$ will force $\mathbf{p}' = -\mathbf{p}$ , but this we still have $p^{0\prime} = +p^0$ after this delta. Thus, the two exponentials now have a residual $\exp(2ip^0x^0/\hbar)$ factor which is new -- fast time dependence!

2) the "spin terms" now have a contribution of the form $i\ (+ p'_0 + p_0\ )\ \sigma^{k0} = 2i\ p_0\ \sigma^{k0}$.

You see both of these results in 3.31 where the cross terms are the last two. The other diagonal term is as simple as the first.

**Fact**: With only positive energy terms, our current **J** in 3.28 was constant in time for our packet. When we add in the negative energy solutions, it suddenly develops ultra-high frequency time dependence from those cross terms! This is the famous **zitterbewebung** of Schrodinger 1930 (jittering agitation, etc). It is an interference effect between the positive and negative energy solutions.

On page 39, we consider a **localized 3D gaussian spatial wavefunction** at t = 0 and ask "does it involve Dirac solutions of both signs of energy as t moves forward? " We assume a gaussian shape and $w^1(0)$ at t=0 as shown in 3.32. We evaluate 3.30 at t=0 to get equation A, something I can verify. We then do a $d^3x$ Fourier on both sides of A. The RHS is done explicitly for us. Authors use a vector $\boldsymbol{\ell}$ as the variable conjugate to **r**, something we usually call **k** instead of $\boldsymbol{\ell}$. The LHS just pins $\boldsymbol{\ell} = \mathbf{p}$ and $-\mathbf{p}$ in the two terms of A, and we get C which I have checked except for the constant. We can then use orthogonality relations to project out the two amplitudes. We find that, although at t=0 we started with a pure u state, we end up t>0 with both u and v states activated, and so we are going to get this jitter effect. The d* amplitude is small in the NR limit since it shows as a v spinor. However, when the gaussian is made small enough that the packet is confined to a **Compton wavelength** λ = h/mc, then p becomes large enough so the negative energy terms are appreciable. The deBroglie is λ = h/p so we are just saying when p gets up near mc, effects start showing up. Of course p can get much larger than mc, since $\mathbf{p} = \gamma\boldsymbol{\beta}mc$ in general.

$\lambda = h/mc = 2.4263102175\times10^{-12}$ meters $\sim 2.4$ pm  Compton, scale of electron extent for QED

As an example of a paradox, B&D consider the famous **Klein Paradox (1929**). The idea is to send in a beam of electrons from the left against a 1D steep potential wall, say a square well wall as shown on page 41. So put an electron in a 1D box, and try to localize it more and more. Make the walls high, that would seem a good idea. We look only at the right wall. We want to get the usual SE exponential damping going into the barrier. We assume the incoming scattering wave is 3.34 first, which is a u(p,s) thing with a small lower part, overall amplitude "a". The reflected wave is assumed a linear combination of w1 and w2, both having small 3,4 parts as shown. The transmitted waves are still free-electron solutions if we shift by the potential wall height $V_0$. So we assume the transmitted wave as shown, and now we have five constants a,b,b',d,d' and we want to solve the problem "as usual" by matching things at the wall. Continuity requires that a+b just to the left of the wall equal d, where we are doing this comparison for the w1 terms. He concludes that b' = d' = 0 due to no spin flip. When you match the lower components at the boundary, you get an a-b condition so you can solve for all the constants, as you would expect. Notice that our interest is in the wave number $k_2$ inside the barrier region. When you increase the barrier high enough, suddenly you get non-exponential transmission! Notice that $j_{trans}/j_{incident} = 4r/(1+r)^2$ where r is the quantity I circled in 3.36. When the barrier gets too high, r goes negative, and we have transmission inside the barrier *going to the left*! And, the reflected current is larger than what came in.

Well, I know that what has happened is we get an electron-positron pair created in addition to the original electron, I guess the positron goes off to the right, making the negative current, and the new electron comes back with the original reflected one and you get larger than came in. In this book, this will be resolved in chapter 5 in terms of "hole theory".

So the Dirac Equation **"knows about"** pair creation from the vacuum. I will be interested to see that chapter.

Meanwhile, there are still many interesting problems to look at where we stay far from this pair-creation energy level. This is where the F-W transformation is going to be useful.

# Chapter 4: The Foldy-Wouthuysen Transformation (46)

**4.1 Introduction (46).** In hydrogen, the Bohr radius is 137 times larger than the Compton wavelength, so we expect good results with predictions we can make which ignore the negative energy solutions. We are not going to be creating many $e^+e^-$ pairs in a cool H atom! The general idea now is to add a potential V(r) to our Dirac theory and see what we can get for solutions say for a central force. The claim is that this 1950 FW work simplifies the problem in a systematic fashion by somehow decoupling our 4-spinor equation into something simpler. BD cannot resist calling FW a "canonical transformation" with no explanation, which leads to the next comment:

Comment on Canonical vs Unitary Transformation: Let U be some transformation that does not involve time. Then we transform the SE in the following manner:

$$H\psi = \hat{E}\psi \quad \Rightarrow \quad HU^{-1}U\psi = \hat{E}\psi \ \Rightarrow \quad UHU^{-1}U\psi = U\hat{E}\psi = \hat{E}\,U\psi \qquad H'\psi' = \hat{E}\,\psi'$$

One example of this kind of thing is moving between "pictures" where $U = e^{iH'}$ where H' is some Hamiltonian. The question now is: in general, must U be unitary? It has to be if we require that the norm of our states be unchanged. If the norms of all states are unchanged, then we will still have total probability = 1 in the primed world if it is so in the unprimed. Note that $\langle\psi'|\psi'\rangle = \langle U\psi \mid U\psi \rangle =$

$\langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\psi\rangle$ only if U is unitary. If U were not unitary, I suppose we could renormalize states in the primed world and make things be OK. But then we would be converting U to unitary in effect.

In Goldstein we read about "canonical transformations" as changing the P and Q of the Hamiltonian perhaps to constant values, and in any event to different values. We did this by *adding* some $\partial F/\partial t$ to the Hamiltonian to form a new Hamiltonian, where F was called a generator. That was classical mechanics. In quantum mechanics, we have our coordinates x and p and our states $\psi$. I don't know right now how you translate the notion of the canonical transformation from classical to quantum mechanics. However, I do know that $U = e^{iHt}$ acts as a unitary transformation which can take your wavefunctions back to time 0 where they are the "initial values", and in classical mechanics you seek a transformation which takes you to a place where P and Q are constants. Maybe somehow the wavefunctions are the coordinates in QM. The web suggests that somehow the "unitary transformation" of QM corresponds to the "canonical transformation" of classical mechanics. For now, this is a semantic question. I suspect this question will get answered if I keep pursuing the general line of inquiry I am now on, I don't need this answer now.

In this chapter, we are going to study various $\psi' = e^{iS}\psi$ transformations which Messiah says are in fact unitary. We will keep finding that $S = -i\beta O/2m$ where $\beta$ is the usual 4x4 matrix and O are some "odd" terms in a Hamiltonian. Thus we have

$$S = -i\beta O/2m \;\Rightarrow\; S^\dagger = +i\, O^\dagger\beta^\dagger/2m = +i\, O\beta/2m = -i\,\beta O/2m = S \;\Rightarrow\; S = S^\dagger \;\Rightarrow\; e^{iS} = \text{unitary}$$

We assumed here the O and $\beta$ anticommute. This is true for example if $O = \boldsymbol{\alpha}\bullet\mathbf{p}$ or $O = \gamma_5$. It is therefore probably true for any possible "odd terms" (defined below) that might occur in the Hamiltonian.

So the FW business is going to involve some unitary transformations which we can "call" canonical transformations if we like. Also, the transformations are "small" in that somehow $S << 1$.

**4.1 Free-particle FW Transformation (46).** Here we are thinking about finding a nice H' in the context of no potentials, just a "free particle".

As a small digression of the authors, consider $H = \sigma_x B_x + \sigma_z B_z$ which is non-diagonal. How would you make it be diagonal? Well, consider $H = \boldsymbol{\sigma}\bullet\mathbf{B}$ in general. We know that

$$R\,\boldsymbol{\sigma}\,R^{-1} = R^{-1}\boldsymbol{\sigma}$$

$$R\,\boldsymbol{\sigma}\bullet\mathbf{B}\,R^{-1} = \mathbf{B}\bullet(R\,\boldsymbol{\sigma}\,R^{-1}) = \mathbf{B}\bullet R^{-1}\boldsymbol{\sigma} = B\bullet\sigma_z\hat{\mathbf{z}} = B_z\sigma_z$$

$$H' = B_z\sigma_z$$

where we have selected R such that $R^{-1}\boldsymbol{\sigma} = \sigma_z\hat{\mathbf{z}}$. In the case they specify, $\boldsymbol{\sigma}$ is in the x-z plane and we just want to bring it back up to the top with a y rotation of a certain angle we could call $\theta$, fine!

Comment: think of R as causing a change in coordinate system in which the problem is studied. When we work with H' as shown above, things are simple. Selecting the z axis to align with B is the best way to work on this problem. The same idea will apply to FW.

In the above, $\sigma_z$ is diagonal of course which is why we like it in H'. In the 4x4 space, we want to get to an H' where all matrices appearing are diagonal at least in the 2x2 subspace sense. Examples include $\gamma_0=\beta$, $\sigma^{ij}$, $\Sigma$. These are called the "even" matrices. Those which are off-diagonal in this sense include $\gamma^i$,

$\gamma_5$ and $\alpha_i$ -- the odd matrices. These "odd" matrices mix the upper and lower components and make our life more difficult.

So, consider our 4D problem. We have

$$H = (\boldsymbol{\alpha}\bullet\mathbf{p} + \beta m)$$

which of course contains those "odd" $\boldsymbol{\alpha}$ matrices we don't like. We seek H' where

$$H' = R(\boldsymbol{\alpha}\bullet\mathbf{p} + \beta m)\, R^{-1} = \text{something with only "even" matrices in it.}$$

$$= p\bullet R\boldsymbol{\alpha}\, R^{-1} + m R\beta\, R^{-1}$$

where now R acts in the 4x4 Dirac space. The suggested purifying rotation is this one:

$$R = \exp([\theta p]\beta\boldsymbol{\alpha}\bullet\hat{\mathbf{p}}) = \exp([\theta p]\boldsymbol{\gamma}\bullet\hat{\mathbf{p}}) = \cos(p\theta) + \boldsymbol{\gamma}\bullet\hat{\mathbf{p}}\,\sin(p\theta)$$

where in the expo expansion we note that $(\boldsymbol{\gamma}\bullet\hat{\mathbf{p}})^2 = 1$ so powers of this thing are either 1 or $\boldsymbol{\gamma}\bullet\hat{\mathbf{p}}$. We then compute $RHR^{-1}$ and we get page 47 G and see the possibility of killing off the term with α's. The killing angle θ is a function of m and p as shown in H. In getting there, we used $\{\alpha,\boldsymbol{\beta}\} = 0$. Repeat: in order to zero out the α term, we choose angle θ as shown. Then using the triangle on the next page to compute the cos and sin in the second term, we get 4.1 which is a very simple result indeed: $H' = \beta\sqrt{m^2 + p^2}$ which is a 4x4 matrix. We have β = diag(1,-1) so the upper and lower component pairs are now decoupled. The eigenstates of H' are going to be four such as (1,0,0,0). The first two have eigenenergy $\sqrt{m^2+ p^2}$, and the last two have eigenenergy $-\sqrt{m^2+ p^2}$. We can ignore the negative energy solutions when we are thinking about the positive energy solutions, even at high velocity, in this representation. Consider back earlier where with the "regular H" we had the four complicated spinors for a particle not at rest. Here we also have four spinors for the same problem, but they all are like (1,0,0,0). So yes, things are much simpler in this "representation" of the hamiltonian.

As an aside note that this rotation thing R is unitary:

$$R = \exp([\theta p]\boldsymbol{\gamma}\bullet\hat{\mathbf{p}}) \quad \Rightarrow \quad R^{-1} = \exp(-[\theta p]\boldsymbol{\gamma}\bullet\hat{\mathbf{p}})$$

$$R^{\dagger} = \exp([\theta p]\boldsymbol{\gamma}^{\dagger}\bullet\hat{\mathbf{p}}) = \exp(-[\theta p]\boldsymbol{\gamma}\bullet\hat{\mathbf{p}}) = R^{-1}$$

**4.2 The FW Transformation with E&M Fields (48).** We now turn on the EM field and try again to "diagonalize" our Dirac equation using some kind of transform, again trying to get rid of "cross terms" or "odd matrices" as they say. In 4.2 we see that the "odd matrices" both involve α, as before.

Now first, jump over to page 49 top where we define a new transformation S such that $\psi' = e^{iS}\psi$. This is not the same S as earlier when we were looking at how spinors transform under Lorentz.

Confusion #1. What happens next in line B is somewhat confusing, so notes are needed. I will do it "step by step" as one always needs to do in BD. On this line we have 4 expressions and 3 equal signs.

(1) the first equal sign with its two expressions is a statement that $H\psi = \hat{E}\psi$ where $\hat{E}$ means the usual $i\partial_t$ operator, so this is the time dependent SE written for $\psi$ with Hamiltonian H -- our starting positions.

(2) The third expression is a rewrite of the second expression, so we understand it and the second =.

(3) the fourth expression is the unexpected one. This comes from the leftmost expression where you apply the time derivative on the two factors. Then it becomes obvious. So now we are happy with B.

(4) Now we take the rightmost equality in B and apply $e^{+iS}$ on the left, and this gives the left equality of C. We then interpret the large square bracket as H' since we then have $H'\psi' = \hat{E}\psi'$. Notice that the $\partial_t$ operator inside H' does NOT act on $\psi'$ (for future reference).

So in all this shuffling, we have simply found the appropriate H' which gives a SE on $\psi'$, nothing more and nothing less. Our goal now is to select some kind of S which will "kill off" our odd matrix terms.

In D we show the usual CH expansion of the sandwich shown in general to all orders. Then in E, only the first four terms of this CH expansion are shown, the fourth being the one with 1/24.

Confusion #2. Now we must understand that claim that S = O(1/m) and what this even means. This seems very unobvious to me because we don't yet know what S is. So now begins some more PhL "interpolative comments" to fill in what BD should have told us but did not to "save space" (our time is not important).

• The first thing we claim is that, in the NR limit, "S is small" so that $\psi'$ is not too different from $\psi$. The idea here is that in the NR limit, the "odd matrix" terms are *already* very small, so it does not take a very large "rotation" to cancel them out. In our statement that $H = \boldsymbol{\alpha}\bullet(\mathbf{p}\text{-}e\mathbf{A}) + \beta m + e\varphi$ [ we have turned on some fields, by the way ] we see that the $\beta mc^2$ term is -- in the NR limit -- very much larger than the other two terms, assuming "reasonable" EM fields. So as just claimed, it is the $\boldsymbol{\alpha}$ term here which is the small odd term we want to get rid of and we will do this by a small rotation S.

• Looking again at H, my first question is this: how does the first term have dimensions of energy? It is really $c\mathbf{p}$ that is sitting there in the first term and this is then $c(\gamma\beta mc) = \gamma\beta mc^2$ [ where $\gamma$ and $\beta$ here are the special relativity things, unrelated to gamma matrix things ]. Then our ratio of interest is roughly just $\gamma\beta$. We set $\gamma = 1$, then as usual our NR limit means $\beta << 1$. This I understand. I do not understand where the ratio 1/m enters the NR limit discussion. As Messiah points out on page 944, the actual smallness ratio is this (which I arbitrarily call $\lambda$)

$$\lambda = [\ \alpha\bullet(p\text{-}eA)]\ /\ m\ =\ [\ c\alpha\bullet(p\text{-}eA)]\ /\ mc^2\ = \text{energy/energy} = \mathcal{J}/m\ \text{ where } \mathcal{J} = \alpha\bullet(p\text{-}eA)$$

Then we see that roughly

$$\lambda\ = \mathcal{J}/m\ \sim pc/mc^2\ = \gamma\beta mc^2/mc^2 = \gamma\beta \sim \beta$$

Now in BD, the thing Messiah calls $\mathcal{J}$ is called $\mathcal{O}$ (for odd terms).

• Conclusion: formally, we have $\lambda = \mathcal{J}/m = \mathcal{O}/m$ as our smallness parameter. For example, we shall see in our first iteration below (BD page 49 F) that $S = (-i/2)(\mathcal{O}/m)\,\beta$ where $\beta$ here is the matrix, so there is the $(\mathcal{O}/m)$ ratio making S be small. We can then "track" our terms by power of (1/m), with the understanding that the actual expansion is in the dimensionless ratio $\mathcal{O}/m = v/c << 1$.

Note in passing: if we have $S = (-i/2)(\mathcal{O}/m)\,\beta$ where $\mathcal{O}$ = odd terms, { in the Hamiltonian formalism we have H(p,q,t) ???} we will have $\partial_t S = (-i/2)\{\partial_t \mathcal{O}\}/m\,\beta$ and this will be non-zero because **A** within $\mathcal{O}$ is in general time dependent. We refer below to $\partial_t S$ as $\dot{S}$, so this shows that in general $\dot{S} \neq 0$ so we have to worry about it.

Now go back to that CH expansion in page 49 D. Tracking with 1/m as discussed above, we see that -- since H is of order m (from βm), and since S is order 1/m, our CH terms are order m, 1, 1/m, $1/m^2$ and so on. Thus, in equation E, the i/6 term is order $1/m^2$ and the 1/24 term is order $1/m^3$ . In this last term, we are allowed to replace H with just βm while maintaining correctness through order $1/m^3$. In other words, this last approximation $H \sim \beta m$ is off by order 1 which would cause a discrepancy of [S,[S,[S, [S, order(1)]]]] which is order $1/m^4$ which we don't care about.

Now looking back at C, we have to worry about $e^{iS}(-i\partial_t)e^{-iS}$ as well as the $e^{iS}He^{-iS}$ . Note that this former thing *cannot* be written as - $(e^{iS}\ \dot{S}\ e^{-iS})$ because, for example, $\partial_t S^2 = S\dot{S} + \dot{S}S \neq 2\,S\dot{S}$. So we have to apply BCH to the operator as shown, and we get this

$$e^{iS}(-i\partial_t)e^{-iS} = (-i\partial_t) + i\,[S, (-i\partial_t)] + i^2/2\,[S,[S, (-i\partial_t)]] + i^3/6\,[S,[S,[S, (-i\partial_t)]]] + ...$$

where I have kept terms through order $S^3$ = order $1/m^3$ . As noted earlier, there is nothing "further to the right" like ψ' that we have to worry about. Thus, we get

$(-i\partial_t)\,1 = 0$

$i\,[S, (-i\partial_t)] = iS\,(-i\partial_t) - i\,(-i\partial_t)S = 0 - \dot{S} = -\dot{S}$

$i^2/2\,[S,[S, (-i\partial_t)]] = i/2\,[S, \{\, i[S, (-i\partial_t)]\,\}] = i/2\,[S, -\dot{S}\,] = -i/2\,[S, \dot{S}\,]$

$i^3/6\,[S,[S,[S, (-i\partial_t)]]] = i/3\,[S,\{\, i^2/2\,[S,[S, (-i\partial_t)]]\,\}\,] = i/3\,[S, -i/2\,[S, \dot{S}\,]\,] = +1/6\,[S, [S, \dot{S}\,]\,]$

And these are then the three terms we see that the end of equation E

Old Comment: The $\partial/\partial t$ term has to be there in C because, if our EM fields are time dependent (which they normally will be except for trivial problems), then S that removes the odd matrices will be a different S for different times, and we have S = S(t). We want H' above to have no odd terms, but B&D claim no S(t) exists which does this for all times t. I will return later to understanding just why that is so.

Old Comment: Meanwhile, looking at page 47 H without the fields turned on, for the non-rel region we will have p/m << 1 and that means tan(2pθ) ~ 2pθ = p/m. This means the S = β**α•p̂** ( pθ) ~ p/m = p/mc to be dimensionless. In non rel, p = mv << mc, so we are to think of S as being order(1/m). In the extreme non-rel limit, we know that coupling goes away between pos and neg energy solutions of Dirac, so not surprising that S is very small in the NR limit, and that S ~ 1/m which I think really means p/m.

Now, in 4.3 we keep terms only through order $m^0 = 1$. So we keep the entire H ( the first term from the $e^{iS}$ H $e^{-iS}$ expansion) plus an order 1 term from the i[S,H] term in the same expansion. Fine. At this order $m^0$ how do we kill off the odd terms? If you select S as shown in F, you get

i[S,β]m = - $\mathcal{O}$ using S = -iβ $\mathcal{O}$/2m easy to show

Notice that $\mathcal{O}$ = **α•**(**p**-e**A**) from page 48 and $\mathcal{E}$ = *e*φ and both are regarded as order $m^0 = 1$, that is, we might have normal sized **A** and φ fields. Therefore, indeed S = -iβ $\mathcal{O}$/2m is order 1/m as claimed earlier. So this tells us that H' = βm + $\mathcal{E}$ for 4.3, and we are diagonal.

Now, using this "first order S", we compute all the other terms in the big mess 49E. The 6 terms of interest are duly computed on page 50, and I am not deriving all these things, nor did I on the last pass (but I could easily do it). We end up finally with this transcription of equation E

H' = βm + $\mathcal{E}'$ + $\mathcal{O}'$ as in 4.4

where we have new odd and even terms. Now the odd terms $\mathcal{O}'$ are order 1/m or higher. Just as we did in the first round, we now define S' = -iβ $\mathcal{O}'$/2m and we know this will knock out our the 1/m portion of $\mathcal{O}'$. So after applying S', we then have

H" = βm + $\mathcal{E}$ ' + $\mathcal{O}$ "

where now $\mathcal{O}$ " is order $1/m^2$. And we do it one more time and we get

H"' = βm + $\mathcal{E}$ ' + $\mathcal{O}$ "' = βm + $\mathcal{E}$ '

where we just throw away $\mathcal{O}$ "' which is order $1/m^3$.

So the point is that we keep doing transformations **to grind down the size of the odd terms in H**. So we end up with what is shown in page 51 A.

Now we do more messy computations to see what these terms are, and we end up with 4.5 where now we see things like **E** and **B** fields along with φ and **A** stuff. There are now lots of terms!

Claim is that the first term is really the square root shown in 51E, where we have the first three terms of this expansion. This is what appeared in 4.1 page 48 with no fields, but we have modified it. But notice that this is not the whole answer! You can't just insert **p**-e**A** into 4.1 and think you are done, that is what has been shown here.

The second term in 4.5 is $-\mu_B\beta$ **σ•B** which I think indicates the normal interaction of an electron with a magnetic field. The moment is the Bohr magneton e/2m. This is what we used in Levitt.

The next two terms are spin-orbit. For central force, only the second of these terms is non-zero and gives the result 4.6. This is famously ½ of what you would expect from a *classical* calculation by saying that the electron sees a **B'** field from the rotating nucleus. The ½ arises because the electron is rotating

around the nucleus and the effect is known as "**the Thomas Precession**". For non-relativistic speeds, this Thomas effect always gives a factor of ½, independent of velocity of rotation. I saved a PDF on this subject.

The very last term in 4.5 is $(e/8m^2)\ \nabla^2 V$ and is called "**the Darwin term**". B&D give a crude explanation of this in 4.7 which is pretty darn close. The idea is that the electron does not see V, but instead sees an effective V because the electron is "spread out" over its Compton size $\lambda = 1/m$.

**Commentary**: I am a little confused about the conclusions of this section. We start with our well-known Dirac H and we go to "some other frame" and observe things there. All the above terms are in the "' frame, with H"'. So I want to say that you are allowed to study your problem in any "frame" or "representation" you want. So here we can choose to work in the H"' representation.

What can I say about equation 4.5 which shows this H"' ? The first "energy term" has $\beta$ and this is mainly what makes the 34 states have negative major energy compared to the 12 states. We also see the $\beta\Sigma$ matrix and the $\Sigma$ matrix by itself appearing in this equation. Note that $\beta\Sigma = \mathrm{diag}(\sigma,-\sigma)$ so those neg energy solutions have backwards magnetic moment in a B field (antiparticles, holes). The spin-orbit terms are the same for 34 as for 12.

This equation has no matrices which couple 12 to 34, and it is "exact" through order $1/m^2$. So, if we want to study 1,2 regular electrons, we can just look at the 12 equation, and that is what 4.5 becomes if you replace $\Sigma$ with $\sigma$ and $\beta$ with 1. The 34 term effects are completely removed. BUT, in order to achieve this, new terms have appeared in the 12 equation!!! These are part of the theory. We have just *exposed* these terms by doing the decoupling.

Now go back to page 12 where we did the non-rel reduction of the Dirac equation in its normal form. In 1.30 we have our two coupled equations. Had we simply set $\chi = 0$, instead of 1.31, we would have gotten 1.34 without the magnetic moment with B term. So this term is really being caused by the small 34 components of the solution being non-zero. I think this is a 1/m effect.

Now about the gyromagnetic ratio thing. On page 13 and with 1.34 we say g = 2. *Why* are we saying this? When you compute the **L•B** term in H either classically or in quantum theory, you get the result being $H' = -\mu_B\ \mathbf{L}\bullet\mathbf{B}$ where $\mu_B = e\hbar/2mc$. The 2 gets there classically because current = ev/(**2**$\pi$r) and area $\pi r^2$ so moment is cM = current x area = evr/2, while L is rmv, so M/L= (evr/2c)/(rmv) = e/2mc. So we could write this as $H' = -g_L\mu_B\ \mathbf{L}\bullet\mathbf{B}$ and say $g_L = 1$. So g = 1 is the "natural thing" to have. But the term in 1.34 showing the **S•B** interaction appears as $H' = -2\ \mu_B\ \mathbf{S}\bullet\mathbf{B}$ so $g_S = 2$. So that is what we are talking about with g = 2.

Now, our reduction on page 12 and earlier takes some account of the small 34 components and out pops this $H' = -2\ \mu_B\ \mathbf{S}\bullet\mathbf{B}$ term. And fiddling with the p•A terms brings out the $-1\ \mu_B\ \mathbf{L}\bullet\mathbf{B}$ term so our Dirac equation is predicting the electrons have g=1 for orbital , but have g=2 for spin. This stuff on page 13 is all quantum, by the way. We got all this stuff just from our simple Dirac reduction.

Now turn to page 51 and H"' as in 4.5. We get the same $H' = -2\ \mu_B\ \mathbf{S}\bullet\mathbf{B}$ term we got on page 12. We also see the exact same term $(\mathbf{p}\text{-}e\mathbf{A})^2/2m$ appearing. But now NEW terms appear which result from the fact that we have been more accurate in accounting for the 34 contributions. We have decoupled through order $1/m^2$ which is stronger than what we did on page 12 which I guess was order 1/m.

So one of the big "new terms" is the famous **"spin-orbit interaction"** ! BD quietly mention this, but the point is this: it is predicted by the Dirac theory!!! I think I have always wondered where the hell it came from in courses I had. As BD point out page 51, it only arises that way with a central potential V(r) which is just fine. Now if V(r) = 1/r, then $1/r\ \partial V/\partial r = -1/r^3$ and we have $H' = -1/r^3\ e/4m^2\ \mathbf{S}\bullet\mathbf{L}$ . If we imagine classically that the electron "sees" a B field $\mathbf{B}' = -\mathbf{v} \times \mathbf{E}$ then our $H' = -2\ \mu_B\ \mathbf{S}\bullet\mathbf{B}$ term would

predict H' = - 2 $\mu_B$ **S•B'**. But crudely L = mvr and vxE = vE = v$\nabla$V = v $\partial$V/$\partial$r, so B' = v $\partial$V/$\partial$r which then tells us H' = - 2 $\mu_B$ **S•** [v $\partial$V/$\partial$r $\hat{\mathbf{z}}$]. But L = mvr so v = L/mr and then
H' = - 2 $\mu_B$/m **S•** [L (1/r)$\partial$V/$\partial$r $\hat{\mathbf{z}}$]. = - 2 $\mu_B$/m **S• L** * (1/r)$\partial$V/$\partial$r.

BUT, the correct result as shown in 4.6 is - 1 $\mu_B$/m **S• L** * (1/r)$\partial$V/$\partial$r which is *half* of our classical prediction for the spin-orbit term. So they say you can think of g = 1 again for this situation, and then both terms involving **L** have this same "orbital g" = 1 (the other L term being on page 13].

But if we think of the electron as being in a non-inertial frame and redo the physics classically using the extra "ω x ..." Goldstein term, then you get the factor of ½. But relativity still comes in because it is what causes the B' field in the first place. I have not figured this all out, but have a PDF on the subject. I cannot digress an infinite number of levels.

The main point is this: The Dirac equation under our FW correctly predicts the spin-orbit term with the correct factor of ½ present. So this just adds to the credibility of the Dirac equation. We needed the FW to bring out these more subtle terms.

The **Darwin term** I have only dimly heard of and I don't recall people worrying much about it. It is a purely radial term for a central potential, so it just shifts energy levels, it does not do radical things like L•S coupling does in terms of state labeling and spectroscopy. I presume we will hear more about it from B&D in next section or two.

**4.3 The Hydrogen Atom (52)**. BD have short-changed this important topic and pretty much state equations 4.9 without any derivation. I think they are claiming that these are the angular parts of a separable solution in spherical coordinates and then four radial functions are installed on page 54 and their equations are given in 4.13 and we are told to go see some old papers given on page 52 bottom. These authors include original papers of Darwin(1928), Gordon(1928), then Bethe & Salpeter (a 1957 book) and Rose (a 1961 book). They then simply quote the energy levels in 4.14.

Do I have the missing meat in some other book? The angular stuff appears not in Schiff or Messiah, but it does appear in my Sakurai, and I am reminded that the latter is well written as I peruse it.

For the moment, let us then accept the eigenenergies as they appear in 4.14. There are two quantum numbers. First is n = 1,2,3.... and second is j which can range only from j+1/2 up to n. For example, if n = 1, I guess you can have j = ½, and if n is 2 you can have j = ½ and 3/2. All very hazy but OK for now. Notice that α = 1/137 appears in 4.14 in two places, actually in the form Zα for hydrogen-like atom. If this is treated as a smallness parameter, you can expand the square root structure in a series and get 4.15 where now the first term with the 1 gives the usual NR energy levels for Z hydrogen. We see there a correction term that is a function of n and j.

Next, it is natural to inquire about the **ground state** and this is n=1 and j= ½ as the only value. The energy is given in the above expansion in equation A, and the actual up and down ground states are given as 4-spinor detailed solutions in B and C. The angular parts are obscure to me and appear to involve $\boldsymbol{\ell}$ = 1 but I know $\boldsymbol{\ell}$=0 in some sense, so we put that on the list of mysteries at this point. The radial behavior lies in factors outside the bracket, including of course the required exponential decay factor. Here they have defined yet another γ as shown top page 56 (a third usage in this chapter really). If Zα > 1, this square root goes imaginary and we get weird Klein-like behavior. But normally we have Zα < 1 and then we may construct a simple table of the lowest states using the historical $\boldsymbol{\ell}$,j,n labels. These four states appear in the page 57 picture along with some other states.

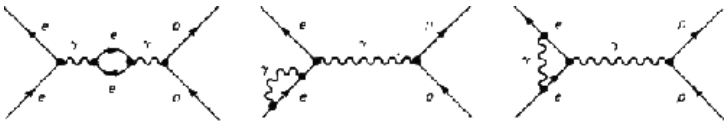
The **hyperfine** stuff is the splitting of all states into doublets due to the nuclear spin coupling with the electron spin in a dipole-dipole interaction a la Levitt. "The 21cm line (1420.4 MHz) was first detected in

1951 by Ewen and Purcell at Harvard University." But that is not really the subject of this chapter, but all lines are shown so doubled in the page 57 figure.

The Dirac theory we are now studying predicts that the that $2S_{1/2}$ and $2P_{1/2}$ should be degenerate (see table energies), but they are not in fact and this is due to the **Lamb shift** which is a field theory effect. On the other hand, Dirac theory does predict a separation between the $2P_{3/2}$ and the $2P_{1/2}$ and this observed split is called the **fine structure** (as opposed to the above hyperfine). This amount is 10.9 GHz, the Lamb shift is 1.057 GHz and the hyperfine is 1.42 GHz.

I now resume the BD batting order:

(p 57) The **Lamb shift** was only discovered in 1947 and BD make only a passing note of it without giving the size. I know it comes from these diagrams, and most of it from the first.




( p 57) The **hyperfine structure** is mentioned.

(p 58) BD then "**estimate**" the size of the hyperfine structure in an obscure way. I think they are saying that the B field is generated by the nuclear spin, but the expression they role out could not possibly be more obscure, they are just showing off here. The reader trying to learn something is never going to know where this comes from, some messy Jackson thing. So this whole page is a complete waste for me. The expression must be doing the relativistic EM field transformation so you get a B field from viewing an electric potential $\varphi = q/r$ from a moving frame, and all that stuff.

(p 58-60, ie through end of the chapter). Here they quote a **simple explanation of the Lamb shift** from some guy Welton. It relies on thinking about the zero point energy of the vacuum. For me, the question is whether Lamb shift comes out of the Dirac theory or not. I think this BD Vol 1 is going to show that the Feynman diagrams are really part of the Dirac theory, for example see page 148, with the hole idea. So I guess my question is this: is there anything you calculate from field theory that contradicts the Dirac theory?

Comment: So I am now disappointed that I invested lots of effort and was not rewarded with a clean derivation of the H atom from Dirac theory. I will have to go read Sakurai or elsewhere for this information. As BD said in their preface, they are not worrying about bound-state problems, so they just fly by this stuff quickly.

I think it would be best for me now to just continue in BD and save the H atom for another time, otherwise I will lose the BD thread and I will have to start over yet again.

## Chapter 5: Hole Theory (64)

**5.1 The Problem of Negative-energy Solutions (64).** An obvious problem with the Dirac spectrum: what keeps positive energy state electrons from "falling" into those negative states? It is easy to compute the transition rate downward and it is infinite. Dirac in 1930 suggested that this would be fixed if ALL those negative energy states were "filled" and that the Pauli Principle applies to those states so you cannot add more, hence no downward transitions.

Pause: exactly "what" is "filling" those negative energy states? Do we suddenly have a theory that is a multi-particle theory and that number of particles is infinite, since that is how many are needed to fill

those negative energy states? Are all those negative energy electrons "really there"? I guess if you were to do Dirac in a Box, you would say that the Box as a whole has all those negative energy electrons in it, not each point in space.

No doubt these questions arise from the fact that this model is not completely correct. Still, it allows lots of predictions. The most obvious is that if you apply a photon with $\geq 2mc^2$ energy (> 1.022 MeV, not very much), you ought to be able to excite one of those negative energy electrons into a positive energy state, leaving behind a positively charged "hole" which might behave like a particle itself. This would sort of test whether those negative energy electrons are "really there". Experimental result: this is exactly what can happen, and the positron was duly found in 1932. The exact date of the time positrons were "postulated" might be 1930 by Dirac instead of 1928 as it says here.

" The existence of positrons was first postulated in 1928 by Paul Dirac as a consequence of the Dirac equation. In 1932, positrons were discovered by Carl D. Anderson, who gave the positron its name.[3] The positron was the first evidence of antimatter and was discovered by passing cosmic rays through a cloud chamber and a lead plate surrounded by a magnet to distinguish the particles by bending differently charged particles in different directions."

So here is the "classic" pattern of particle physics. A new particle type is predicted, then it is found. So Dirac discovered anti-matter, pretty amazing.

**5.2 Charge Conjugation (66)** The "absence" of a negatively charged of energy -E might be the same as the "presence" of a positively charged particle of energy +E. Now things get a little tricky and I want to at least attempt to follow the logical steps here. Our regular Dirac equation is 5.1. If there were a positive energy positron, it's equation would be 5.2 where the sign of e is changed, all else is the same. Here notice authors have used $\psi_c$ thinking of the positron as a "charge-conjugated electron".

So what operator will generate $\psi_c$ from $\psi$? The answer is going to be this: $\psi_c = (i\gamma^2)\psi^*$.

Note added: You need to change the relative sign of just the $\not{A}$ term in 5.1. We know $\gamma^2 \not{a}^* \gamma^2 = \not{a}$ and $\gamma^2 1\ \gamma^2 = -1$ so if you do a K sandwich then a $\gamma^2$ sandwich, the $i\not{\nabla}$ term negates, the m term negates, and the $\not{A}$ term stays the same, mission accomplished.

In the development of this result, authors show that the * makes the grad term go negative, leaving you with 5.3, Then the matrix $i\gamma^2$ negates $\gamma^{*\mu}$ and causes the relative change in the sign of the mass term, and thus the combination gives you 5.2 from 5.2 as desired.

Now BD manage to confuse things a bit with their notation. They know it is convenient to write this result as $\psi_c = (i\gamma^2)\psi^* = (i\gamma^2\gamma^0)\gamma^0\psi^* = C\,(\gamma^0\psi^*)$ where they define $C = i\gamma^2\gamma^0$. The reason for defining C in this way is this: (note that $\bar{\psi}^T$ is a column vector, though at first you think it is a row vector)

$$\gamma^0\psi^* = \bar{\psi}^T = (\psi^\dagger\gamma^0)^T = \gamma^0\,(\psi^\dagger)^T = \gamma^0\,(\psi^*)$$

which allows us to say that

$$\psi_c = C\bar{\psi}^T \qquad C = i\gamma^2\gamma^0 = i\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\begin{pmatrix} 0 & -\sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} = -i\,\alpha_2$$

$$C^T = -C \qquad C^* = C \qquad C^\dagger = -C \qquad C^2 = -1 \qquad C^{-1} = -C$$

$$C^{-1}\,\gamma^{\mu}\,C \;= -\,(\gamma^{\mu})^{T}$$

where the phase "i" is chosen to make the matrix elements of C be real. The matrix $i\gamma^2$ is written out in 5.7 and you see them there computing $\psi_c = (i\gamma^2)\psi^*$ starting with $\psi$ being the spin-down negative energy state r=4 shown in A. As shown, you end up with the spin-up state r =1. So "the absence of a negative energy spin-down electron" equals "the presence of a positive energy spin-up positron". So this is the first mention of what happens to "the spin" in charge conjugation. I guess it has to be that way. You remove something with $J_z = -1/2$ and this leaves the sea with a net $J_z = +1/2$.

On scratch, I have shown that if you take the general p spinors shown on page 30, the result is what you would expect. You find this:

$$(i\gamma^2)[\,w^4(\mathbf{p})\exp(-i\varepsilon_4 p^{\mu}s_{\mu}/\hbar)]^* = w^1(\mathbf{p})\exp(-i\varepsilon_1 p^{\mu}s_{\mu}/\hbar) \qquad (i\gamma^2) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

where I have used the fact that $\varepsilon_4 = -1$ and $\varepsilon_1 = +1$. From the phase part, we would say that the r=4 solution had $E = p^0$ and the resulting r=1 solution also has $E = p^0$, in this tricky ε notation. Also, we can see that the $(i\gamma^2)\psi^*$ operation "does not change **p**". That is to say, the $w^4$ spinor starts off with $p_-$ and $-p_z$, but the matrix changes to $+p_z$ and the * to $p_+$ which are what you find in $w^1$. We defined $w^4(\mathbf{p})$ as a function of **p** in the way it appears, so we cannot argue with the resulting claim that **p** is the same.

Thus, in the sense just described, the entire $p^{\mu}$ is *the same* for both $\psi$ and $\psi_c$ ! Again, you might think of $\psi$ as being negative energy and having momentum **-p**, and its absence is $\psi_c$ which has positive energy and positive **p**.

So in the above analysis, we have to keep careful track of the ε factor in order to reach our conclusion that $p^{\mu}$ is the same for $\psi$ and for $\psi_c$. This is reflected in the energy projector part of 5.8 where you have to think of $\varepsilon < 0$ for the first projector acting on $\psi$, and $\varepsilon > 0$ for the projector in the last line which acts on a positive energy state.

A similar interpretive thing happens with the $s^{\mu}$ vector. In the rest frame, I have shown on scratch that

$$\tfrac{1}{2}\,(1 + \gamma_5 \not{s}) = \tfrac{1}{2}\,(1 + \gamma_5\gamma^{\mu}s_{\mu}) \;=\; \tfrac{1}{2}\,(1 - \gamma_5\gamma^{i}s^{i}) \;= \tfrac{1}{2}\,(1 + \gamma_0\boldsymbol{\Sigma}\bullet\mathbf{s})$$

which agrees with their slightly erroneous mid-page claim on page 69 (they have the final paren in the wrong place, a rare BD typo). Here **Σ** is the extended **σ** matrix. To show the above, you just multiply out the matrices in the larger 2x2 sense. In that world, $\gamma_5$ is a "reverser", etc. So, since $\gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ we would say that our starting spinor $w^4$, although "spin down" with **-s**, picks up the -1 from the $\gamma_0$ matrix in the projector, whereas the $w^4$ spinor with **+s** gets the +1 from the $\gamma_0$ matrix, so we have a double negative in the first case. This then is why the covariant projector $\tfrac{1}{2}\,(1 + \gamma_5 \not{s})$ is saying that both states have the same $s^{\mu}$. Again, we can interpret this in the "absence/presence" mindset if we want: the $w^1$ state has positive spin because it is the absence of the negative energy state which has negative spin.

This then is why we get the results shown on page 69 A. Remember that $w^4$ (not $w^3$) was v(p,s) and $w^1$ was u(p,s) and we see that there are no minus signs. So now we see why these u and v objects were

defined as they were, so that they come out to be "charge conjugates" of each other! Perhaps later authors will fiddle with the arbitrary phase factor shown.

Technical Note concerning 5.8. Notice the object $\not{p}^T = p_\mu{}^T \gamma^{\mu T}$ . What is meant by $p_\mu{}^T$ ? Normally that would change a column vector to a row vector, but I think they really define $\not{p}^T$ to mean $p_\mu \gamma^{\mu T}$ so only the matrix is transposed and $p_\mu$ are just some coefficients. One of my $\gamma$ rules is that $\gamma_o \gamma^{\mu\dagger} \gamma_o = \gamma^\mu$ so if we add the effect of the *, then we can say that $[\gamma_o \not{p} \gamma_o]^* = \not{p}^T$, so this explains how these T objects are appearing in the second line of 5.8. To get to the last line, you have to insert $C^{-1}C$ twice and recall from page 67 the result $C^{-1}\gamma^\mu C = -\gamma^{\mu T}$ and that clinches the fact that the first terms both get sign-changed and we then end up with the last line of 5.8.

Authors state clearly that $\psi_c$ is the positron wavefunction.

In mid page 69 authors describe a more general operator $\mathcal{C}$ in an operational sense which I would say is a symmetry operation of the Hamiltonian or Lagrangian. You do *, then $C\gamma^0$ on spinors, then negate the entire $A_\mu$ potential. I am quite familiar with this kind of symmetry, and the point made here is that if $|\psi>$ is a Hilbert space state for an electron, then there must also exists states $\mathcal{C}|\psi>$ for the anti-particle, so we are definitely predicting the existence of positrons.

**5.3 Vacuum Polarization (70)** Now we have to think of an isolated normal electron as being "surrounded" by the Dirac Sea of negative energy and negatively charged electrons (not positrons!). Does the Coulomb force couple between positive and negative energy electrons? If so, we might expect the sea electrons to be "repelled" away from the isolated pe electron causing a region around our electron to have an effective positive charge distribution as illustrated in the second figure, where the first shows the charge distribution in some sense for the electron. The region of repelled sea electrons seems similar to the way a charge in a dielectric would align electric dipoles, causing a region of opposite polarization charge around the charge in question. For this reason, this effect is called "vacuum polarization" meaning Dirac Sea polarization. The figures indicate some scale R for this polarization, but nothing is said about R. Perhaps the "bare charge" is really 2e, but when you test from far away, you see the "dressed charge" which is only e, with the Dirac sea providing -e>0 from its integral. These are very vague notions that get supported later in field theory. There, if you are very close with your "test charge", energies are high and you can make pairs and these are associated with the "vacuum polarization" Feynman diagram and this is in fact the major part of the Lamb shift (which seems to contradict their comment here). But in field theory, "polarization" means a slightly different thing. You actually create a particle pair, causing a charge separation which is then the "polarization". Again, the Dirac Sea model is not quite physical.

They "sidestep" the final question of why we don't see all that negative charge of the Sea with the comment that if it made an E field, what direction could it point in? Sure, an infinite uniform charge distribution can't make an E field.

**5.4 Time Reversal and Other Symmetries (71)** First, we are reminded that $P = \gamma^0$ times a phase, and how this expression for P causes the Dirac equation to have form invariance, we did this back in Sec 2.3.

Note added: What happens to the spin projector with P?

$$\gamma^0 \tfrac{1}{2}(1 + \gamma_5 \not{s}) \gamma^0 = \tfrac{1}{2}(1 + [\gamma^0\gamma_5\gamma^0][\gamma^0 \not{s} \gamma^0]) = \tfrac{1}{2}(1 - \gamma_5 s_\mu [\gamma^0\gamma^\mu\gamma^0])$$

$= \frac{1}{2}(1 - \gamma_5 s_\mu \gamma^{\mu\dagger}) = \frac{1}{2}(1 - \gamma_5 s_\mu \gamma_\mu) = \frac{1}{2}(1 - \gamma_5 s^\mu \gamma^\mu) = \frac{1}{2}(1 + \gamma_5 \not{s}')$ where $s'_\mu = -s^\mu$.

which says that $s'^0 = -s^0$ but $\mathbf{s}' = \mathbf{s}$, so "the spin" in the rest frame does not change. This seems correct because we think that P converts $w^1(\mathbf{p})$ to $w^1(-\mathbf{p})$ so both should be "spin up". What about the phasor factor"

$p^\mu s_\mu \rightarrow (p_\mu)(-s^\mu) = -p^\mu s_\mu$ so s.p must be a pseudoscalar.

This seems correct because parity should change p but not an angular momentum thing which is a pseudovector. That is: vector • pseudovector = pseudoscalar.

Our problem now is to make the time-dependent SE have *form invariance* under $t \rightarrow t' = -t$. BD kind of bollix this up a bit, I redid it on scratch. You start with 5.11, replace $\psi = \mathcal{T}^{-1}\psi'$, apply $\tau$ from the left and you get this result for 5.12:

$-(\mathcal{T}\, i\, \mathcal{T}^{-1})\, \partial_{t'}\psi' = \mathcal{T}\, H\, \mathcal{T}^{-1}\psi'$

Trial and error shows that you should make $\mathcal{T} = TK$, where K means complex conjugation. Then the above becomes

$+ i\, \partial_{t'}\psi = T\, H^*\, T^{-1}\psi' = T\,[\, \boldsymbol{\alpha}^* \bullet (+i\boldsymbol{\nabla} - e\mathbf{A}(t)) + \beta m + e\varphi(t)\,]\, T^{-1}\psi'$

where we know that $\boldsymbol{\nabla}$ and e and $\mathbf{A}$ and m and $\beta$ and $\varphi$ are all real. Then we rewrite, time-reversing the potentials and thinking of a current making $\mathbf{A}$, and also priming $\nabla$, to get

$+ \partial_{t'}\psi = T\, H^*\, T^{-1}\psi' = T\,[\, \boldsymbol{\alpha}^* \bullet (+i\boldsymbol{\nabla}' + e\mathbf{A}(t')) + \beta m + e\varphi(t')\,]\, T^{-1}\psi'$

So we will achieve form invariance with the original equation $i\partial_t\psi = H\psi$ *if* the following properties are true:

$T\, \boldsymbol{\alpha}^*\, T^{-1} = -\boldsymbol{\alpha}$ and $T\, \beta\, T^{-1} = \beta$

On scratch I showed that a solution to this problem is this

$T = i\gamma^1\gamma^3 = T^{-1}$ (it turns out) $\mathcal{T} = TK$

where as usual we pick a phase arbitrarily. Recall the rules for "sliding $\gamma^i$ around": (1) sign change every time it passes through some other gamma of the set 0,1,2,3 that is different. (2) get minus when one kills itself as in $\gamma^j\gamma^j = -1$. So once again, BD take something simple and make it unclear enough that I have to go off and do it myself.

In 5.16 we now look at the effect of $\mathcal{T}$ on an E>0 electron state which has $p^\mu$ and $s^\mu$ . The $\mathcal{T}$ sandwich around a real 4-vector slash item like $\not{p}$ results in $\not{p}'$ where the spatial part changes sign. So this sort of

reinforces the idea that time reversal negates both **p** and **s** as you would expect "if time went backwards". The buzz phrase for this is "**Wigner time reversal**".

Now suddenly we start talking about something called $\psi_{PCT}(x'^{\mu} = -x^{\mu})$. First, let's just assemble the pieces to compute such an item.

$$\psi_{PCT}(x'^{\mu} = -x^{\mu}) \;=\; \mathcal{PCT}\psi(-x) = \mathcal{P}\,(\mathrm{CK}(\mathrm{TK}(\psi\text{-}t, -x))) = \mathcal{P}\,(\mathrm{CK}(\mathrm{T}(\psi^*(t, -x))) = \mathrm{P}(\mathrm{C}(\mathrm{T}^*(\psi(t, x)))$$

$$= e^{i\varphi\gamma_0}\; i\,\gamma^2\; (-i\gamma^1\gamma^3)\,\psi(t, x) = e^{i\varphi}\,\gamma^0\gamma^2\,\gamma^1\gamma^3 \;= -\,e^{i\varphi}\,\gamma^0\,\gamma^1\,\gamma^2\gamma^3 \;= i\,i\,e^{i\varphi}\,\gamma^0\,\gamma^1\,\gamma^2\gamma^3$$

$$= i\,e^{i\varphi}\,\gamma_5\,\psi(x)$$

which agrees with what they got. Without question, if the thing on the right were going backwards in space and time, then the thing on the left is going forwards in space and time. So now let's also assume that the thing on the right has E < 0, ie, is negative energy and is a momentum eigenstate hence $\varepsilon < 1$ and has some spin vector $s^{\mu}$. We can then play with the projectors as shown in 5.18. We start with an $\varepsilon<1$ electron with p and s as shown. After fiddling, we find the $\psi_{PCT}$ has $\varepsilon>0$ with the same p and reversed s. Recall that from 5.8 that C alone gave you the same p and s, so the new thing here is that s got negated. So here is the bottom line: "positron $\psi$ with E>0 [ namely, $\psi_{PCT}(x)$] = electron $\psi$ with E<0 multiplied by $ie^{i\varphi}\gamma_5$ and going backwards in spacetime [ namely $ie^{i\varphi}\gamma_5\psi(-x)$ ]. "

Now with the fields on, this claim is confirmed as we move from A to B on page 74. To get from one to the other, you need to do the usual sandwich trick but this time with the PCT operator $ie^{i\varphi}\gamma_5$. As usual, this is non-trivial so let's write it out:

$$H = \boldsymbol{\alpha} \bullet (-i\nabla - e\mathbf{A}) + \beta m + e\varphi \qquad \text{with } H\psi = -\,E\,\psi \qquad E > 0 \quad \text{equation A}$$

If we sandwich PCT around this H, the complex conjugations I think cancel out from C and T, and we are left worrying about [ note that $\gamma_5\boldsymbol{\alpha}\,\gamma_5 = +\,\boldsymbol{\alpha}$ since both $\gamma^0$ and $\gamma_i$ get sign changes ]

$$\gamma_5[\boldsymbol{\alpha} \bullet (-i\nabla - e\mathbf{A}) + \beta m + e\varphi]\,\gamma_5 \;= [\,+\,\boldsymbol{\alpha} \bullet (-i\nabla - e\mathbf{A}) - \beta m + e\varphi\,]$$

where we have only worried about the $\gamma_5$ effect. Next, we need to negate spacetime in all the arguments. The potentials don't change so we have

$$= [\,+\,\alpha \bullet (+i\nabla' - e\mathbf{A}') - \beta m + e\varphi'\,] \;= -\,[\,\boldsymbol{\alpha} \bullet (-i\nabla' + e\mathbf{A}') + \beta m - e\varphi'\,]$$

and this is the exact negative of what we started with except for spacetime reversing primes and the charge sign being reversed. . Then we move the sign to the right and end up with

$$[\,\boldsymbol{\alpha} \bullet (-i\nabla' + e\mathbf{A}') + \beta m - e\varphi'\,]\,\psi'_{PCT} \;= +E\,\psi'_{PCT} \qquad \text{// this is equation B}$$

Were we to change the sign of e, we recover the same form as the original equation.

I will now stick with the previous quote and say:

"a positron ψ with E>0 and q = |e| [ namely, $\psi_{PCT}(x)$] = electron ψ with charge -|e| and E<0 multiplied by $ie^{i\varphi\gamma_5}$ and going backwards in spacetime [ namely $ie^{i\varphi\gamma_5}\psi(-x)$ ]. "

Since we know what the charges are of positrons and electrons, we can omit that part and rephrase as:

"positron ψ with E>0 [ namely, $\psi_{PCT}(x)$] = electron ψ with E<0 multiplied by $ie^{i\varphi\gamma_5}$ and going backwards in spacetime [ namely $ie^{i\varphi\gamma_5}\psi(-x)$ ]. "

I don't think you can leave out the " multiplied by $ie^{i\varphi\gamma_5}$" part of this statement.

This concept of interpreting a positron as related to an electron going backwards in spacetime is called "the **Feynman-Stuckelberg** form or positron theory." paper 1941.

The last paragraph of this section and the chapter is another comment on the notion of the **"symmetries"** of the Hamiltonian or Lagrangian, hence "of the theory" ( I should say, symmetries of the equations of motion, I suppose). You can consider different "interaction terms" in the Hamiltonian and ponder what symmetries they have. The feeling is that the Dirac theory *ought* to apply to "other" spin-1/2 particles like protons such that all three of our basic symmetries P,C, and T are preserved.

Having said this, authors make two closing comments:

(1) Lee and Yang showed that the weak interaction violates both P and C. The present theory of weak interactions and experiment confirm this fact (my comment).

(2) Authors claim that PCT combined must always be a symmetry if you assume proper Lorentz invariance and the usual spin-statistics relationship. But they do not prove this claim! I am not sure where I have a proof of this theorem (but I surely do), and not clear how to find this on the web.

My impression is that PCT combined is what gives you the true antiparticle.

# Chapter 6: Propagator Theory (78)

**6.1 Introduction.** The aim is "to calculate things" in principle "exactly". Field theory is a big load, so they want to put it off. Follow the Feynman method.

### 6.2 The Nonrelativistic Propagator (78)

Quick Math Background on Green's Functions. Stakgold Volume 1 page 261 defines what is meant by a Green's Function. You have a differential equation Lu = f(x) let's say in 1D, where L is a *differential* operator. Then the Green's Function g(x',x) = G(x'-x) is the solution to the differential equation when the RHS is "driven by" a delta function, so we have Lg(x,x') = δ(x-x'). The usual method is to first solve a problem for the Green's function with a δ driving source, such that the Green's meets the boundary conditions. Then, once you have the Green's, you can write down the solution to Lu = f(x) as follows:

$$u(x) = \int dx'\, g(x',x)\, f(x')$$

The reason being that

$$Lu(x) = \int dx' \, Lg(x',x) \, f(x') = \int dx' \, \delta(x-x') \, f(x') = f(x)$$

This is the basic idea. We can work with a differential equation Lu = f(x) and try to solve it for u(x) with some boundary conditions, or we can find the Green's function from Lg(x,x') = δ(x-x') also subject to some boundary conditions, and then our solution is $u(x) = \int dx' \, g(x',x) \, f(x')$. I recall doing this a lot in Jackson's book and class where the differential equation was perhaps the Poisson equation $\nabla^2\varphi = \rho$ and we found Green's functions that met certain boundary conditions like constant on a sphere, say. I cannot remember whether there were scattering situations in Jackson that used Green's.

In our current context, we have a *partial* differential equation and multiple variables. Fine.

In this chapter, the Green's Functions will be "particle propagators", amplitude to go from x to x'.

Application of the above. In this section we will use the following Green's functions:

$$(i\partial_t - H) \; G(x,x') = \delta^4(x-x')$$
$$(i\partial_t - H_o) \; G_o(x,x') = \delta^4(x-x')$$

We can then show that "Huygens's Principle" of 6.1 is just a restatement of the SE. There we write the Huygen's rule as this

$$\theta(t'-t) \, \psi(x') = i \int d^3x \, G(x',x) \, \psi(x) \qquad \text{// we only scatter forward in time! Causality.}$$

If we apply $(i\partial_t - H)$ to the LHS, we get 0 (from the SE) + the term generated when $i\partial_t$ hits the θ function which is just iδ(t-t') ψ. Meanwhile, on the RHS $(i\partial_t - H) \; G(x,x') = \delta^4(x-x')$ and the $\delta^3$ kills off the integral, and the δ(t-t') is left over and we get the identity that iψ = iψ. That is why the i has to be on the RHS, because it is part of $i\partial_t$.

So *my* main point here is this: We don't have to "accept" Huygen's Principle 6.1 as some kind of non-derivable starting point for this section, thank goodness. In fact, we just derived 6.1 above, and we find that it is entirely equivalent to the SE; it has the same information, it is an alternate form of the SE. The claim will be that this form is more conducive to having scattering type boundary conditions.

Now we can start the fancy presentation of this section. In 6.2 we "turn on" a potential for differential short $\Delta t_1$ at time $t_1$. In 6.3 we integrate 6.2 over time from $t_1$ to $t_1+\Delta t_1$ and Δψ includes only the *additional* amount of Δψ caused by the presence of V ≠ 0 during this time, so we don't show the base component of Δψ which we get integrating 6.2. We have φ (plane wave) on the right because that's what ψ was before V hit at time $t_1$,.

Then in 6.4 we insert this special Δψ amount on the right of a contribution form of 6.1, the i's cancel. After the tiny $\Delta t_1$ scattering event, we propagate with no potential hence with $G_o$. So the claim is we are showing a particular "contribution" lying within 6.1 which is this momentary scattering contribution of V turned on for just a moment at time t1. Turn V off all the time, this contribution is zero.

Then 6.5 tells us that the total amplitude at x' is the original wave φ plus this scattered contribution. This explains the first equality in 6.5. Here is a derivation of the second equality, where we use some compacted notation:

$f = x' = x_{final}$

$a = x = x_a.$

$$\psi_f = \varphi_f + \int d^3x_1\, G^o_{f,1} V_1 \varphi_1\, dt_1 \qquad \varphi_1 = i\int d^3x_a\, G^0_{1,a}\, \varphi_a$$

$$\Rightarrow \quad \psi_f = i\int d^3x_a\, G^0_{f,a}\, \varphi_a + \int d^3x_1\, G^o_{f,1} V_1 [i\int d^3x_a\, G^0_{1,a}\, \varphi_a]\, dt_1$$

$$= i\int d^3x_a \{\ G^0_{f,a} + \int d^3x_1\, dt_1\ G^o_{f,1} V_1\, G^0_{1,a}\ \}\, \varphi_a$$

These two terms appear as diagrams (a) and (b) on page 80. We can regard the quantity in {...} as our full G(f,a) for this particular situation, and that is what 6.6 says.

Next we turn on V again at $t_2 > t_1$ . This causes *two* new terms, the single scattering of (c) and the double scattering of (d) and these terms are written in 6.7.

And so we build up a mountain of scattering terms of different orders. In 6.9 = 6.10 authors show all single, double and triple scattering terms in compressed notation with the time ordering maintained as shown. If we put a time θ function inside $G_o$ ( we then have a retarded propagator with > speed of light) we get 6.11. If you break out just the very last scattering, you get the **integral equation 6.12** where we now have some assumptions about convergence of the series and V being small, etc. Of course that will be the case in QED which is mostly what this book is about with the 1/137 factor. We can think of our drawings as "Feynman diagrams" for our current example which is the non-relativistic SE since we have no 4x4 matrix stuff going on (although $H_o$ has not yet been stated).

Now we come to the notion of a scattering experiment and boundary conditions. We assumed some initial state φ that was "far from" the action of the potential V which we now assume is somehow localized near a scattering center of some sort. So we already have that BC worked into our equations above. Equation 6.13 shows the full G carrying the propagation load from the distant past to some present spacetime point x'. This is an $id^3x$ thing like 6.1. Equation 6.14 inserts 6.12 into 6.13 to get the first expression, but then it uses another version of 6.13 going to point $x_1$ to replace the $id^3x\ G(1...)\varphi(x..)$ with ψ(1) and this gives the second expression in 6.14. So this thing is now a traditional "integral equation" for the wavefunction ψ, and it involves $G_0$ and V as shown. I recall integral equations which "show the last rung" from my multiperipheral model work.

Things are clarified in the next "Note Added".

---

Note added: Let's look at these things in matrix form a bit:

Notation: I denote $d^3x$ integrals with a ".", and $d^4x$ integrals with no "."

We have: (I am also keeping track of the factors or i ). When φ appears on the right, it usually means we have taken it into the far past. If only in the finite past, it would be ψ such as in 6.1

$\psi' = iG.\psi$ 6.1 or 6.17, ψ in the finite past , this is Huygens' = SE

$G = G_0 + G_0VG$ 6.12 integral equation for G, driven by V and $G_o$

$\psi' = iG.\varphi$ 6.13 prime means x' and t' in future
$\varphi' = iG_0.\varphi$ how $G_0$ would move $\varphi$ into the future

$\psi' = iG.\varphi = i(G_0 + G_0VG).\varphi = (iG_0.\varphi) + G_0V(iG.\varphi) = \varphi' + G_0V\psi$ , so we have

$\psi' = \varphi' + G_0V\psi$ 6.14 integral equation for $\psi$, driven by V and $G_o$

We need some comments here!! In the above, $\varphi$ always refers to a simple plane wave or state of system where V = 0. If you apply the full complexity of scattering represented by G, you get at a later time that $\psi' = G.\varphi$, so symbol $\psi$ here represents some highly complicated wavefunction that is, if you like, the superposition of an infinite number of Feynman diagrams. But if we use only the straight-through diagram which means $G_0$, we get $\varphi' = G_0.\varphi$ as the evolution. In the above equations, I had to prime things just to show them as being in the future to avoid writing things like $\varphi = G_0\varphi$ which, nevertheless, is a valid matrix equation! It says $\varphi(x') = \text{Int}_x G_o(x',x)\varphi(x)$ or $\varphi_{x'} = (G_0)_{x',x}\,\varphi_x$. So we have two integral equations above with are really equivalent:

| | | | |
|---|---|---|---|
| $G = G_0 + G_0VG$ | 6.12 | integral equation for messy G | $d^4x$ |
| $\psi = \varphi + G_0V\psi$ | 6.14 | integral equation for messy $\psi$ | $d^4x$ |

What about expressions for the S-matrix? In 6.16 we see this:

$$S_{fi} = \varphi_f^{\dagger}.\psi_i^{(+)} = \varphi_f^{\dagger}.\, iG.\varphi_i = \varphi_f^{\dagger}.\,(\varphi_i + G_0V\psi_i^{(+)}) = \varphi_f^{\dagger}.\,\varphi_i + \varphi_f^{\dagger}.G_0V\psi_i^{(+)}$$

$$= \delta_{fi} + \varphi_f^{\dagger}.G_0V\psi_i^{(+)}$$

This is written in this odd way because the second term has all graphs with 1 or more V interactions, so we have exposed the "straight through" term as $\delta_{fi}$. We could replace $\psi_i^{(+)} = iG.\varphi_i$ on the right to get this alternative form

$$S_{fi} = \delta_{fi} + i\,\varphi_f^{\dagger}.G_0VG.\varphi_i$$

And here for fun is a longer derivation of these last results:

$$S_{fi} = \varphi'^{\dagger}_f.\psi_i^{(+)} = \varphi'^{\dagger}_f\,.\,(iG.\varphi_i) = i\varphi'^{\dagger}_f.G.\varphi_i$$

$$= \varphi'^{\dagger}_f.\,(iG_0 + iG_0VG)\,.\varphi_i = \varphi'_f\,.\,(iG_0\,.\varphi_i) + i\varphi'^{\dagger}_f\,.G_0VG.\,\varphi_i = \varphi'_f\,.\varphi'_i + i\,\varphi'^{\dagger}_f\,.G_0VG\,.\varphi_i$$

$$= \delta_{fi} + i\,\varphi_f^{\dagger}\,.G_0VG\,.\varphi_i$$

So we can summarize all these ways of writing $S_{fi}$

$S_{fi} = \delta_{fi} + \varphi_f^{\dagger}.G_0V\psi_i^{(+)} = \delta_{fi} + i\,\varphi_f^{\dagger}.G_0VG.\varphi_i = \varphi_f.G.\varphi_i$

and 6.16 is the first equality above.

Now let's do *one more* form for $S_{fi}$. Suppose we write out $G_o$ as the basis function sum shown in 6.26, then the term shown above becomes

$\varphi_f'^{\dagger}.G_0V\psi_i^{(+)} = \varphi_f'^{\dagger}.\{-i\,\theta \int d^3p\, \varphi_p'\, \varphi_p^{\dagger}\}\, V\psi_i^{(+)} = -i\,\varphi_f^{\dagger}\, V\, \psi_i^{(+)}$ for going into future

Notice that we have used orthogonality to say $\varphi_f^{\dagger}.\,\varphi_p = \delta^3(\mathbf{p}_f - \mathbf{p}) = \int d^3x \langle \mathbf{p}_f | x't'\rangle\langle x't'|\mathbf{p}_i\rangle$ (the dot means a $d^3x$ integral). I have put a few primes on things in the above equation line to show that the orthogonality is allowed because the two factors in question are at the same space and time. In the final result above, notice there are no "dots" now. We have therefore a single $d^4x$ integral. So we can add this "one more form" to our list above and summarize as follows:

The plane waves in the above are really $\varphi_p = e^{-ip\bullet x}\, e^{-iEt}/(2\pi)^{3/2}$ which then gives us the delta function as claimed.

**S-Matrix Summary**

There are very many ways to write $S_{fi}$ and they are all "derived" right here. Form 5 was derived about 2" up above. Form 1 is really the definition of the S matrix. Supporting material comes first.

| | | | |
|---|---|---|---|
| $\theta\psi' = iG.\varphi$ | 6.13 | prime means x' and t' in future | (0) |
| $\theta\varphi' = iG_0.\varphi$ | how $G_0$ would move $\varphi$ into the future | | (1) |
| $G = G_0 + G_0VG$ | 6.12 | integral equation for messy G | (2) |
| $\psi = \varphi + G_0V\psi$ | 6.14 | integral equation for messy $\psi$ | (3) |

| | | | |
|---|---|---|---|
| $S_{fi} = \varphi_f^{\dagger}.\psi_i^{(+)}$ | form 1 | 6.16 | defines $S_{fi}$ |
| $S_{fi} = i\,\varphi_f^{\dagger}.G.\varphi_i$ | form 2 | 6.30 | f1 + (0) |
| $S_{fi} = \delta_{fi} + i\,\varphi_f^{\dagger}.G_0VG.\varphi_i$ | form 3 | | f2 + (2) + (1) |
| $S_{fi} = \delta_{fi} + \varphi_f^{\dagger}.G_0V\psi_i^{(+)}$ | form 4 | 6.16 | f1 + (3) |
| $S_{fi} = \delta_{fi} - i\,\varphi_f^{\dagger}\, V\, \psi_i^{(+)}$ | form 5 | 6.34 | see note above |
| $S_{fi} = \delta_{fi} + \varphi_f^{\dagger}\, V\, G.\varphi_i$ | form 6 | 6.33a | f5 + (0) |
| $S_{fi} = \delta_{fi} - i\,\varphi_f^{\dagger}\, V\, (\varphi_i + G_0V(\varphi + G_0V\psi))$ | form 7 | | f5 + (3) twice |
| $= \delta_{fi} - i\,\varphi_f^{\dagger}\, V\, \varphi_i - i\,\varphi_f^{\dagger}\, V\, G_0V\, \varphi_i - i\,\varphi_f^{\dagger}\, V\, G_0V\, G_0V\, \varphi_i + \ldots$ | | | 6.33b |

This last form *seems* to suggest that (recall $1 + x + x^2 + ... = 1/(1-x)$ )

$$iG = 1 - iV - iV G_0 V - iVG_0VG_0V + ... = 1 - iV(1 + G_oV + G_0V G_0V + ...) = 1 - iV(1/[1- G_0V])$$

$$\Rightarrow iG - 1 = -iV(1/[1- G_0V]) \Rightarrow (iG - 1)(1- G_0V) = -iV \qquad ????$$

If V = 0, the last result seems to say that $iG_o = 1$ which is true sort of in the sense of (1) above. Well, BD don't really do justice to a formal presentation of "scattering theory" but I know I have other books that do. So we have to put this in the same category with the hydrogen atom solution. At least the equations above with BD equation numbers must be true.

One could imagine all these things written in bra-ket operator notation. Then G would be the S matrix operator! It controls the full amplitude between the $\varphi_i$ initial state and the $\varphi_f$ final state.

**Comment on the i's.** I said earlier that iG and $iG_o$ are natural combinations. Another natural combination is -iV, which I failed to notice. This appears for example in (6.3) on page 79, and this gets there due to the "i" in 6.2 which is $i\partial_t$. This means that the combinations VG and $VG_0$ don't change sign if you "show the i's", namely, VG = (-iV)(iG). So wherever these two appear together, there is no "i" to think about. So the integral equation $G = G_0 + G_0VG$ can be thought of as $iG = iG_0 + iG_0(VG)$ . This idea gives you a way to check the signs in all the above equations. When the i's are properly bound, all signs are +. For example, we will always have $S_{fi} = \delta_{fi}$ + something. The first -order perturbation result is a good example. It says $S_{fi} = \delta_{fi} + \varphi_f^\dagger (-iV) \varphi_i$ .

---

On page 83 we then assume an initial φ of a plane wave, our final observation is a similar plane wave at some different $k_f$ wavevector and we define the S matrix $S_{fi}$ as in 6.16 right at the start. The pass through of the original wave gives the usual k-space delta function, and the second term contains the trailing part of 6.14 still with the full ψ in there, now called $\psi^+$ because it is moving forward in time from the staring position as $\varphi_i$. We can then use the diagrams to make a series of terms for our S matrix amplitude. This is a little rough right now, but this is exactly the path we will take.

This entire discussion appears as Chapter 9 in Schiff which is probably my best reference. I wonder if I have some notes on that Schiff chapter? My "scattering" sources and notes are very wide spread!

**Comments**: This S matrix business on page 83 applies to the scattering of a single particle by a scattering center, think of Rutherford scattering of alpha particles by a nucleus with no recoil. The single particle in question has wavefunction $\psi^+(x,t)$. The scattering center is described by some V(x,t). So at this point, we don't have an S matrix with several particle legs coming in and several coming out, which is more what I am used to. So in this one-particle case, we have the incoming and outgoing asymptotic plane waves φ to consider and that is about it.

**6.3 Formal Definitions and Properties of Green's Functions (still non-relativistic situation)**

Stakgold has much to say about Green's functions, in fact Chapter 1 is so titled. Here we are dealing with an "equation of evolution" with a linear time derivative which is like the heat = diffusion equation. The logic on page 84 is just showing that our Huygen's Principle now written as 6.17 is equivalent to the

Green's function *definition* that $LG = \delta^4$ which is then stated in 6.22. Along the way we use the interesting integral representation of the δ and θ functions shown on page 84. The time θ (Heaviside) turns out to be absolutely crucial to everything, you cannot haphazardly ignore it! I will just restate what I said above

$(i\partial_t - H)\ G(x,x') = \delta^4(x-x')$
$(i\partial_t - H_o)\ G_o(x,x') = \delta^4(x-x')$

$\theta(t'-t)\ \psi(x') = \int d^3x\ G(x',x)\ \psi(x)$ // we only scatter forward in time! Causality.

Starting **page 85**, authors undertake to actually compute $G_o(x,x')$, something I recall doing long ago somewhere. The first step is to get the problem over into Fourier momentum space (p = k) and we quickly find the result 6.25 that says $G_o(p,\omega) = 1/(\omega - p^2/2m + i\varepsilon)$. By choosing the sign of the ε in this way, in 6.26 we are then able to generate the correctly signed θ function in the resulting $G_o(x,x')$. Here, the resultant $G_o$ in spacetime (not momentum space) in 6.26 is written as a spectral sum over our plane wave basis functions that looks a lot like the completeness condition.

We see that our plane wave situation is just a special case of a more general situation. Assume that you have a differential operator L in space and time, like $(i\partial_t - H)$, where we are in an "evolution situation". If we assume that G can be written as in 6.28, which is a θ times a spectral sum of the eigenfunctions of L, then if we apply L to both sides look what happens: the LHS of course just becomes $\delta^4(x-x')$ . The RHS has a 0 term where L hits on one of the ψ's (having the right x argument), but also the term you get when $\partial_t$ hits on the theta function. This creates δ(t-t') which forces the time arguments in the sum to be the same, and they *then* produce $\delta^3(\mathbf{x}-\mathbf{x}')$ as per the completeness condition in 6.27. So we then get LHS = RHS = $\delta^4(x-x')$, so our assumption that G can be written as in 6.28 must be true. In the footnote, authors state the closed-form result of integrating the RHS of 6.26 and they comment that this is very similar to a diffusion equation result they know of.

Now backing up, the $\psi_n$ are eigenfunctions of $L\psi_n = \lambda_n \psi_n$ which is $H\ \psi_n = E_n\ \psi_n$ in our case. In general we have these being orthonormal and this fact is then time-independent as I now show:

$$< \psi_n(t)|\ \psi_m(t) > = <U\ \psi_n(0)|\ U\psi_m(0) > = < \psi_n(0)|\ U^\dagger\ U\ \psi_m(0) > = < \psi_n(0)|\ \psi_m(0) > = \delta_{mn}$$

where $U = \exp(-iHt/\hbar)$ is the usual Hamiltonian time driver which is unitary because H is Hermitian. This same idea of time independence of the inner product is used in the next paragraph!

We now come to **page 87** top. Authors first show that the same G which takes ψ forward in time, propagates ψ* backwards in time. You have to stare at things a little to see where x and x' are located in 6.29 versus the previous equation.

The rest of this section seems to rehash the scattering expansion for G, and for the S matrix S. I don't think there is anything new here. Then at the end, authors show that **S is unitary** because the orthonormality of the initial plane waves is maintained as they go through the scattering region. This is the famous "unitarity of the S matrix" which I know just says the sum of the probability of all outcomes is 1. Geoff Chew always thought this would act as a serious constraint on the hadronic scattering amplitude perhaps enough so to actually *determine it* along with things like analyticity and crossing. I think he has given up on that program but of course one can never tell.

Equation 6.30 is still a *single-particle* S-matrix amplitude and you have to keep thinking of the scattering center as a potential V perhaps distributed over some space near **r** = 0 and time near t=0. A

plane wave $\varphi_i$ arrives from the distant past, and we look at the various outgoing plane waves $\varphi_f$ in the distant future.

Equation 6.31 breaks out the H' term of the Ham and puts it on the right, a no brainer. The second line is just fiddling with delta functions, nothing new here. But then as I have marked in the book, you can treat the common delta function on the RHS as $LG_0$ acting with the x' variable, and you can extract the operator portion L out of the x" integral. Then you compare the things acted upon by this operator and set them equal. This is loosely referred to as "doing an integral", and we then get $G = G_0 + G_0VG$ where I am now using a matrix form to account for spacetime integrations, and this is then 6.32 which is just an integral equation for G whose kernel is $K = G_0V$. We could identify this integral equation with page 195 Stakgold equation 3.6 where we set u = G, k = VG, $\mu = 1$, and $f(x) = G_0$. We think either of just one variable of G to get u. This is a Fredholm integral equation. You can see how this integral equation leads to a sum of series situation,

$$G = G_0 + G_0VG = G_0 + G_0V(G_0 + G_0VG) = G_0 + G_0V(G_0 + G_0V(G_0 + G_0VG)) \text{ etc}$$

$$= G_0 + G_0VG_0 + G_0VG_0VG_0 + G_0VG_0VG_0VG_0 + ...$$

We saw this thing on page 81 in 6.9 earlier and rewritten in 6.11 with built-in time thetas and then 6.12 was our integral equation $G = G_0 + G_0VG$. So this is why I say we are doing a "rehash" here of what we already did before.

If we insert the above expansion for G into our S matrix formula 6.30 we get a series for $S_{fi}$ which reflects this series, and a typical term in this series would be shown in the page 90 picture where, for now, we restrict attention to going sequentially forward in time with our scattering interactions with Vat different spacetime locations, still presumably glommed near x=0 and t=0.

**6.4 The Propagator in Positron Theory (89).** The title of this section is a little odd, I think it should be The Propagator in the Dirac Theory of the Electron and Positron. The previous section lays the groundwork where we deal with full spacetime $d^4x$ stuff and we have the "retarded" Green's function propagator with its internal $\theta$ function which maintains causality. In this upcoming long 10-page section, we will have to add all the 4x4 matrix machinery, and we will have to deal with the negative energy solutions as positrons and all that stuff, it should be an interesting voyage.

The opening 3 pages through bottom page 92 are considering the same scattering scenario as in the non-rel previous section, but with a few differences. We no longer distinguish time, but imagine a scattering center to be $V(x)d^4x$. Then we get to the pictures on page 91 which we knew were coming. We are told on page 90 that "one may say that at each vertex, a particle is destroyed and another created", anticipating Volume II and QFT. Then on page 92 we are told we can interpret the backwards scattered paths as electrons going backwards in spacetime which we otherwise interpret as positrons going forward in spacetime. Then we can have arbitrary zigzagging of the lines always shown forward on page 90. They are "relying on intuition" rather than "rigor".

So finally on bottom page 92 we start in on the 4x4 gritty details. Recall in the previous section we defined our Green's in 6.22 by applying $L = (i\partial_t - \hat{H})$ and getting $LG = \delta^4$. I put a hat on the H to emphasize that it is a spatial differential operator that acts on G(x,x'). So the analogous thing in the Dirac theory is to take $L = (\hat{\not{p}} - e\not{A} - m)$ which again is a spatial differential operator, and it is also a 4x4 matrix. We then state this matrix equation $LG = \delta^4$ and the full G is called $S'_F$. We are of course allowed to ask

for the G of any operator L we come up with, and $S'_F$ is the Green's for L as just stated, no matter that we have $L = \gamma^0(i\partial_t - H)$. Authors are putting primes on everything to indicate x' in the operator which is fine, I might have switched x and x' in this presentation. The prime on $S'_F$ is how they want to indicate the "full" propagator, reserving the symbol $S_F$ for the free propagator. That is, $G = S_F'$ and $G_0 = S_F$ when we compare to the non-rel earlier presentation.

So of course 6.40 then defines $S_F$. We then repeat the earlier trick of going to momentum space and we quickly determine that $S_F(p)$ is as shown in 6.42 where the famous $1/(\not{p}-m)$ is just a shorthand notation for the item shown, and I note in passing that $\not{p}\not{p} = p^2$ which is easy to show and which BD have not mentioned ever.

Now when we try to return from p-space to x-space, we have our $d^4p$ integration to carry out and, as before, we ask about the contour for the $dp^0$ integration. The integrand has poles in $p^0$ at the locations shown in the drawing, with the displacements also as shown. These displacements are positioned for a special reason. Based on the form $S_F(x' - x)$ going from x to x', if t'>t we are moving into the future, we need to close the contour down to get expo damping $\exp(-ip^0(-iR)) = \exp(-p^0R)$. This picks up the pole on the right side which is the "positive energy pole". We are going the "wrong way" around this pole so that makes one minus sign. The pole residue is then $-2\pi i$ (1/2E) and we get the form shown in 6.44 where BD have written out the $\not{p}+m$ numerator in detail showing the positive energy E. So when you go into the future, this scheme says you only have positive energy contributions in the integral which makes $S_F$. Looking at 6.43 where we see an example of a positive energy positron $\psi_c$ and a positive energy electron $\psi$, we note that both these things have positive frequency. Charge conjugation you may recall was $\psi_c = (i\gamma^2)\psi^*$, so although we started with $\psi$ as a negative energy $\psi$, the * makes it positive energy.

So the point of this is that our first form 6.44 includes only the effect of positive energy positrons and positive energy electrons moving into the future.

We then obtain 6.45 for "going into the past" and we find all negative energy components, again, with our pole position decision. It does seem anti-causal to be able to propagate anything into the past, but we always have this interpretation issue. Also, a LT can change the meaning of "the past". Look at page 91 first figure. We can think of pair production as a pe positron and a pe electron going into the future. OR, we can think of starting at the upper left and we have a positron going into the past and then it turns into a positron going into the future. I think we are going to need "both directions" to get things to work right. By the way, in getting 6.45 we had the normal pole contour direction, so you might think there should be an opposite sign. But the residue is now 1/(2*-E) and that makes another minus sign, so these two minus signs cancel and 6.45 then has the same minus sign that 6.44 has.

In any event, we see what the pole position choice does, and I agree with 6.44 and 6.45. These two equations are combined in 6.47 using time theta functions and projectors for the $\not{p}+m$ type factors. Notice that 6.46 is somewhat out of presentation order and has $d^4p$. They are just showing how a single $i\varepsilon$ in the denominator can cause all the correct pole positions. The (m/E) factor arises from the projector definition as shown (the m) and from the pole residue shown in 6.44 and 6.45 (the E).

Now for the first time in the book, in equation A they write the properly normalized plane wave including its: normalization factor, the spinor $w^r$, and the spacetime expo (which has $\varepsilon_r$ in it). We can then verify 6.48 by pulling off the expo and normalization stuff, and using the facts claimed top of page and earlier in my notes regarding the alternate forms for the projectors, see discussion above about equation 3.9b and "completeness". Again, this was something BD did not mention anywhere. So I am very happy with 6.48.

What about 6.49? My first hour spent on this had me trying to insert 6.48 for $S_F$ into the RHS of 6.49. I kept making two bad mistakes and I was getting bogus results.

The first mistake is this: the spinors show up in the form $ww^\dagger$ in 6.48, which is really a 4x4 matrix (a completeness partial unity). The equation they quote as "with the aid of" is 3.11 on page 31, but this is of the form $w^\dagger w$ which is a scalar. So it is totally unclear to me how I could make use of 3.11 in the process just described. Notice that the actual component indices of the spinors are not written anywhere on this page. My mistake was accidentally using the scalar 3.11 result for the $ww^\dagger$ matrix object.

My second mistake in the above program was thinking I could do the $d^3p$ on the spatial expo to get $\delta(\mathbf{x}-\mathbf{x}')$. This was wrong because **p** also shows up in the energy part of the expo in the form $\sqrt{p^2+m^2}$, so you don't really have a simple integral representation for a $\delta^3$ function!

I don't really know how they intended to use 3.11 to verify 6.49, but I have my own way. Just apply the operator $(\hat{\not{p}}' - m)$ to both sides using result 6.40. Since $\psi^+$ satisfies the Dirac SE, on the LHS we pick up a term only where we trip up on the theta function. That is to say

$$(\hat{\not{p}}' - m) = \gamma^0(i\partial_{t'} - \hat{H}') \text{ and so } \gamma^0(i\partial_{t'})\,\theta(t'-t) = +i\gamma^0\,\delta(t'-t).$$

So action with $(\hat{\not{p}}' - m)$ on the LHS of 6.49 gives $+i\gamma^0\,\delta(t'-t)\psi^{(+)}(x')$.

Now apply $(\hat{\not{p}}' - m)$ to the RHS. We get the $\delta^4$ as in 6.40 and RHS $= i\gamma_0\,\psi^{(+)}(x')\,\delta(t'-t)$ and we have then shown that both sides match under the action of this operator. This is good enough proof for me. Notice that the only use of the (+) label is that this reminds us we are propagating into the future and that sets the sign of the argument of the $\theta$ function on the LHS. And notice that the $\gamma^0$ must be there on the RHS of 6.49 as they show, because my proof needs it there!

The derivation of 6.50 is exactly the same, but the (-) label goes with the negated $\theta$ function argument, and of course that causes a minus sign if you redo the line of algebra shown above.

I am sure I could find a better way to "prove" these guys 6.49 and 6.50. I recall a certain trick but not clearly enough to use it. It lets you do the $d^3p$ integration even though that energy expo is there. You elevate to a $d^4p$ integration by adding another delta function, then something happens and you solve the problem. I could not find this trick in my binder TK notes, but I am sure it will arise sooner or later in my readings.

So we are now at the bottom of page 95. The Swiss Ernst Stuckelberg figured things out in 1942 but didn't do much with it, Feynman put the thing together in 1948 and gets most of the credit because he actually did calculations and got right answers. The name Dyson is not mentioned by authors.

Now top **page 96** repeats our previous gimmick of breaking out the H' term from the Ham and putting it on the RHS. Now this term is $e\not{A}S_F'$ rather than VG. We then get our integral equation for the full $S_F'$ as we got before, now we have the 4x4 matrix stuff as well, but we still write: (see note added above)

$$G = G_0 + G_0VG \quad (6.12) \qquad \rightarrow \qquad S_F' = S_F + S_F\,e\not{A}\,S_F' \qquad // \text{ 6.51}$$

and we have our "alternative" integral equation mentioned above

$$\psi = \varphi + G_0V\psi \quad (6.14) \qquad \rightarrow \qquad \Psi = \psi + S_F\,e\not{A}\,\Psi \qquad // \text{ 6.53}$$

In this last case, we were already using the symbol $\psi$ for the plane waves, so we had to go to $\Psi$ for the full "messy" wavefunction (spinor).

Now what are the next two equations 6.54 and 6.55 saying? The messy object $\Psi$ exists at all times. If we evaluate $\Psi(x,t)$ when $t \to -\infty$, it is so far back in the past that the only possible contributions to it from the integral $\Psi - \psi = S_F e\not{A} \Psi$ must be " backwards" contributions with negative energy propagation and these come from the part of $S_F$ which has the negative energy projector $\Lambda_-$ in 6.48 so we get only the $1^-$ factor which is the r=3,4 sum, and we have then 6.55. The reverse statement is true to get 6.54.

Now comes the mystery comment in the paragraph after 6.55. The first comment is about 6.54. We are talking forward scattering here of an electron. We have the $\Lambda_+$ projector sitting there. So we need now more interpretation of 6.54, having already derived it. In the rightmost integral, we have the incoming messy electron $\Psi$ scattering only into positive energy plane wave states $\bar{\psi}^r_p$ for r = 1,2. This integral all by itself is some kind of S scattering amplitude. Of course p is summed over $d^3p$. We then take this amplitude integral and multiply it by $\psi^r_p$ now at x which is the final location to get our final $\Psi(x)$. So somehow this is saying that if we decompose what is happening inside the integral into a momentum integral of plane wave intermediate states, we find that only positive energy states are present. Sort of a filter. Hence the comment that the "electron" $\Psi$ cannot "fall into" negative energy states in this integral after scattering by $\not{A}$. OK, these are just interesting words and phrases, the equations have the truth.

Now suddenly we have 6.56 but this is just a translation of our earlier non-rel formulas. Recall from above that we had this as one of our ways to write $S_{fi}$

$$S_{fi} = \delta_{fi} - i \varphi_f^\dagger V \psi_i^{(+)} \qquad \text{form 5}$$

=>

$$S_{fi} = \delta_{fi} - i \bar{\psi}_f e\not{A} \Psi_i^{(+)} \qquad // \ 6.56$$

where again we make the changes $V \to e\not{A}$, $\psi \to \Psi$, $\varphi \to \psi$, $^\dagger \to$ overbar. The only complication is that we have a different sign in 6.56 if the final state is a negative energy state, and this arises from the minus sign in the $1^-$ projector noted above somewhere.

Now to page 97 of this action-packed drama. We have our S-matrix amplitude in 6.56, and we can insert into this the integral equation 6.53 "iteratively" (as we have done earlier in these notes) and we get 6.57 for the $n^{th}$ order term. There are n powers of e here, and n occurrences of $\not{A}$ and n integrals. This is a Feynman graph with n vertices and we know that all time orderings are allowed and in fact contribute. That is to say, we would include a path like that on page 90 as well as a path like that in 6.5 b.

---

Very Long Note inserted:

The purpose here is to obtain *all* the Dirac results which correspond to all our non-rel results. Many of these items don't get any mention in this chapter. Again, a problem with the haphazard presentation method of our authors. But I am happy to do it. This note goes on for about 7 pages and ends with our a summary table.

Let's compare this with our earlier non-rel perturbation expansion

$S_{fi} = \delta_{fi} - i\,\varphi_f^{\dagger}\, V\, \varphi_i - i\,\varphi_f^{\dagger}\, V\, G_0 V\, \varphi_i - i\,\varphi_f^{\dagger}\, V\, G_0 V\, G_0 V\, \varphi_i + \ldots$ **6.33b**

$S_{fi} = \delta_{fi} - i\,\bar{\psi}_f\, e\not{A}\, \psi_i - i\,\bar{\psi}_f\, e\not{A}\, S_F\, e\not{A}\, \psi_i - i\,\bar{\psi}_f\, e\not{A}\, S_F\, e\not{A}\, S_F\, e\not{A}\, \psi_i + \ldots$ **6.57**

Where does the $\bar{\psi}$ bar come from? Let's review a bit the parallel development in previous and present sections. We have

$(i\partial_t - H_o)\, G_o(x,x') = \delta^4(x-x')$ $\quad H_o = \hat{p}^2/2m$ for free particle
$(i\partial_t - H)\, G(x,x') = \delta^4(x-x')$ $\quad H = H_0 + V$

$\gamma^0(i\partial_t - H_o)\, S_F(x,x') = \delta^4(x-x')$ $\quad H_o = \boldsymbol{\alpha} \bullet \hat{\mathbf{p}} + \beta m = \gamma^0(\boldsymbol{\gamma} \bullet \hat{\mathbf{p}} + m)$ for free particle
$(\not{\hat{p}} - m)\, S_F(x,x') = \delta^4(x-x')$

$\gamma^0(i\partial_t - H)\, S_F'(x,x') = \delta^4(x-x')$ $\quad H = H_0 + V = H_0 + e\varphi - \boldsymbol{\alpha} \bullet (e\mathbf{A})$ // see 4.2 p48
$(\not{p} - e\not{A} - m)\, S_F'(x,x') = \delta^4(x-x')$ $\quad = \gamma^0(\boldsymbol{\gamma} \bullet \hat{\mathbf{p}} + m + e\,\gamma^0\varphi - \boldsymbol{\gamma} \bullet (e\mathbf{A})) = \gamma^0(\boldsymbol{\gamma} \bullet \hat{\mathbf{p}} + m + e\not{A})$

So the first thing we notice (moving toward an answer to our question) is that an extra $\gamma^0$ appears on the LHS in the Dirac case compared to the non-rel case, and we see that this was put there so that $S_F$ comes out being a world-scalar, or at least so the defining equation has a covariant look: $(\not{\hat{p}} - m)\, S_F = \delta^4$. So now we have a $\gamma^\circ$ "floating around".

What is the correponding Huygens' Principle?

$\theta(t'-t)\varphi' = iG_0.\varphi$ how $G_0$ would move $\varphi$ into the future // see 6.1 with $G_o$ not G
$\theta(t'-t)\psi' = iS_F\,\gamma^\circ.\varphi$ how $S_F$ would move $\psi$ into the future // see 6.49

$\theta(t'-t)\psi' = iG.\psi$ the "full" versions of the above
$\theta(t'-t)\,\Psi' = iS_F'\,\gamma^\circ.\,\psi$

Recall from above that $(\not{\hat{p}}' - m) = \gamma^\circ(i\partial_{t'} - \hat{H})$ and so $\gamma^\circ(i\partial_{t'})\,\theta(t'-t) = +i\,\gamma^\circ\,\delta(t'-t)$ which says we get $+i\,\gamma^\circ\,\delta(t'-t)\psi$ for the LHS of the second line above. The RHS then uses $(\not{\hat{p}} - m)\, S_F(x,x') = \delta^4$ and we get RHS $= i\,\delta(t'-t)\,\gamma^\circ.\psi$. This then is WHY the $\gamma^0$ appears in the Huygens for $S_F$. It has to be there due to the way $S_F$ is defined. So we see that $\gamma^0$ continues to "float into" our scattering equations.

Now in the Dirac world, we have a Huygens' for going backwards in time. It is the same as forward except for a minus sign, so

$\theta(t-t')\,\psi' = -\,iS_F\,\gamma^\circ.\psi$

You verify this the same was as before, and now $(i\partial_{t'})$picks up a minus sign on the left, which then matches the mins sign on the right. So perhaps we would combine these together and write

$\theta\,\Psi' = \varepsilon\, iS_F'\,\gamma^\circ.\,\psi$ (0) where $\varepsilon$ is the energy signature of $\psi$, and $\theta$ as apropo

What about all the other non-rel equations, how do they appear in Dirac world? We have already seen above (derived BD page 96) that the integral equation for $S_F$ is exactly parallel to that for G,

$G = G_0 + G_0VG$ (6.12) $\rightarrow$ $S_F' = S_F + S_F e\not{A} S_F'$ // 6.51

so there are no $\gamma^\circ$ factors in this integral equation for $S_F'$. But let's review anyway how this works:

non-rel:

$$(i\partial_{t'} - H_{o'})G = \delta^4(x'-x) + VG = \int d^4x'' \{\delta^4(x'-x)\} [\delta^4(x''-x) + V(x'') G(x'';x) ]$$

$$= \int d^4x''\{(i\partial_{t'} - H_{o'})G_o\} [\delta^4(x''-x) + V(x'') G(x'';x) ] =$$

$$= (i\partial_{t'} - H_{o'}) \int d^4x'' G_o [\delta^4(x''-x) + V(x'') G(x'';x) ]$$

$\Rightarrow$ $G = G_o + G_oVG$

relativistic

$$(\hat{p}' - m)S_F' = \delta^4(x'-x) + e\not{A} S_F' = \int d^4x'' \{\delta^4(x'-x)\} [\delta^4(x''-x) + e\not{A}(x'') S_F'(x'';x) ]$$

$$= \int d^4x''\{(\hat{p}' - m)S_F\} [\delta^4(x''-x) + e\not{A}(x'') S_F'(x'';x) ] =$$

$$= (\hat{p}' - m) \int d^4x'' S_F [\delta^4(x''-x) + e\not{A}(x'') S_F'(x'';x) ]$$

$\Rightarrow$ $S_F' = S_F + S_F e\not{A} S_F'$

So this shows that no $\gamma^0$ appear here because nothing we do "exposes" a $\gamma^0$.

Next, let's examine item (3) in our table above which says $\psi = \varphi + G_0V\psi$ in the non-rel case. Let's repeat the above comparision to see what happens with this integral equation for the wave function.

non-rel: $\psi' = iG.\varphi = i(G_0 + G_0VG)\varphi = ( i G_0 \varphi) + i G_0V(i G\varphi) = \varphi' + iG_0V\psi$

rel: $\Psi' = i \varepsilon S_F'\gamma^\circ. \psi = i \varepsilon (S_F + S_Fe\not{A}S_F') \gamma^\circ. \psi = (i \varepsilon S_F\gamma^\circ.\psi ) + S_F e\not{A} (i \varepsilon S_F'\gamma^\circ \psi) = \psi' + S_F e\not{A} \Psi$

So here we see the $\gamma^\circ$ making an appearance, but in the end they disappear. Also, $\varepsilon$ appears, but then it too goes away. The above equation appears as 6.53,

So now we have formed the start of the long table we had for the non-rel case, to wit:

| | | |
|---|---|---|
| $\theta \Psi' = \varepsilon iS_F' \gamma^\circ. \psi$ | (0) | forward in time $\theta(t'-t)$ and $\varepsilon = 1$, else both opposite |
| $\theta \psi' = \varepsilon iS_F \gamma^\circ.\psi$ | (1) | |
| $S_F' = S_F + S_F e\not{A} S_F'$ | (2) | |

$$\Psi' = \psi' + S_F\, e\not{A}\, \Psi \qquad (3)$$

Now finally we come to the S matrix stuff. The definition would be this:

$$S_{fi} = \varphi_f^{\dagger}.\psi_i^{(+)} \quad \Rightarrow \quad S_{fi} = \psi_f^{\dagger}.\,\Psi_i^{(+)} \qquad \text{form 1}$$

where the $^{\dagger}$ is there this reason: The non-Dirac scalar product is $\langle\psi|\Psi\rangle = \langle\psi|x\rangle\langle x|\Psi\rangle = \psi^{\dagger}\,\Psi$ where if we want to use x-space matrix notation, $\psi^{\dagger} = \psi^{*T}$ is a row vector of starred items, the star from the Hilbert space basic inner product definition in QM. In the Dirac case, we have this same notation, but the $^{\dagger}$ has the additional meaning of transpose in the 4x4 Dirac matrix space.

Now we can install our Huygen's into the above to get

$$S_{fi} = \psi_f^{\dagger}.\,\Psi' = \psi_f^{\dagger}.\,(iS_F'\,\gamma^{o}.\,\psi_i) = i\,\psi_f^{\dagger}.S_F'\,\gamma^{o}.\,\psi_i \qquad \text{form 2} \qquad \text{// analog } i\,\varphi_f^{\dagger}.G.\varphi_i$$

$$S_{fi} = \delta_{fi} + i\,\varphi_f^{\dagger}.\,S_F\, e\not{A}\, S_F'.\,\psi_i \qquad \text{form 3} \qquad \text{f2 + (2) + (1)}$$

$$S_{fi} = \delta_{fi} + \psi_f^{\dagger}.\,S_F\, e\not{A}\, \Psi \qquad \text{form 4} \qquad \text{f1 + (3)}$$

Now we come to "form 5" which we know has a fancier derivation than the other forms. We are going to start with the term above $\psi_f^{\dagger}.\,S_F\, e\not{A}\, \Psi$ and insert $S_F$ from 6.48 , (lower line is for comparison only)

$$\psi_f^{\dagger\prime}.\,S_F\, e\not{A}\, \Psi = \psi_f^{\dagger\prime}.\{-i\theta \int d^3p\, \Sigma_{r=1,2}\psi_p^{r\prime}\,\bar{\psi}_p^{r} + i\theta_b \int d^3p\, \Sigma_{r=3,4}\psi_p^{r\prime}\,\bar{\psi}_p^{r}\}\, e\not{A}\, \Psi$$

$$\varphi_f'^{\dagger}.G_0 V\psi_i^{(+)} = \varphi_f'^{\dagger}.\{-i\,\theta \int d^3p\, \varphi_p'\,\varphi_p^{\dagger}\}\, V\psi_i^{(+)} = -i\,\varphi_f^{\dagger}\, V\,\psi_i^{(+)} \qquad \text{for going into future}$$

As before, the primes indicate $x'^{\mu}$ coordinates space and time. How does $\bar{\psi}$ suddenly appear. It appears because in the Dirac world, completeness and orthogonality have bars, as shown page 30, and this allows for covariant equations. The bars above come from our 1+ positive energy completeness = the positive energy projector.

Now in the Dirac world, we are using the normalization shown in page 95 A. The factor $\sqrt{m/E}$) was not present in our non-rel work, nor of course was the spinor, though the $\pi$ factor was present.

Now in the non-rel case, we just asumed $\varphi_f$ was a plane wave with $\mathbf{p}_f$ momentum and the spatial integration yielded up a $\delta^3(\mathbf{p}_f - \mathbf{p})$. In the Dirac case, in order to continue we need to specify whether $\psi_f$ is a pe or a ne plane wave with $p_f$. So let's first assume that $\psi_f = \psi_f^{s}$ = plane wave with s = 1 or 2. In this case, according to 3.11 we get no contribution from the second term, and get a hit on the r = s term in the first term, so we get

$$\psi^{s}{}_f^{\dagger\prime}.\,S_F\, e\not{A}\, \Psi = \psi^{s}{}_f^{\dagger\prime}.\{-i\theta \int d^3p\, \Sigma_{r=1,2}\psi_p^{r\prime}\bar{\psi}_p^{r}\}\, e\not{A}\, \Psi$$

$$= -i\theta \int d^3p\, \psi^{s}{}_f^{\dagger}.\psi_p^{s}\ (\bar{\psi}_p^{s} e\not{A}\, \Psi)$$

where we have again used 3.11 to remove the r=1,2 sum (it hits on one term or the other). At this point, the implicit $d^3x'$ of the "." integration does this:

$$\int d^3p \quad \int d^3x \sqrt{m/E'}\sqrt{m/E} \; * w^{s\dagger}(\mathbf{p_f})\, w^s(\mathbf{p})\, e^{+ipf.x'}\, e^{-ipi.x'}/(2\pi)^3 =$$

$$\int d^3p \quad \sqrt{m/E'}\sqrt{m/E} \; * w^{s\dagger}(\mathbf{p_f})\, w^s(\mathbf{p}) \int d^3x\, e^{+ipf.x'}\, e^{-ipi.x'}/(2\pi)^3 =$$

$$\int d^3p \quad \sqrt{m/E'}\sqrt{m/E} \; * w^{s\dagger}(\mathbf{p_f})\, w^s(\mathbf{p})\, \delta^3(\mathbf{p_f} - \mathbf{p_i})\, \text{expos_time} = \quad (m/E)\, w^{s\dagger}(\mathbf{p_f})\, w^s(\mathbf{p_f}) = 1$$

where in the very last step we have used 3.11 for the spinor contraction! Thus, we have shown that, in this positive energy final state case, we have

$$\psi^s{}_f{}^{\dagger\prime}.\, S_F\, e\not{A}\, \Psi \quad = \psi^s{}_f{}^{\dagger\prime}.\{ -i\theta \int d^3p\, \Sigma_{r=1,2} \psi_p{}^{r\prime} \bar{\psi}_p{}^r \}\, e\not{A}\, \Psi \quad = -i\theta\; \bar{\psi}_p{}^s\, e\not{A}\, \Psi$$

and so we get our "form 5" S-matrix expression,

$$S_{fi} \;=\; \delta_{fi} \;+ \psi_f{}^{\dagger}.\, S_F\, e\not{A}\, \Psi \quad = \delta_{fi} \;- i\, \bar{\psi}_f{}^s\, e\not{A}\, \Psi \qquad\qquad \text{form 5}$$

where our final probing state is the spinor plane wave with s = 1 or 2. We could write I suppose

$$\bar{\psi}_f{}^s \;= (\sqrt{m/E})\, (1/2\pi)^{3/2}\, e^{+ipf.x}\; \bar{w}^s(p_f) \qquad\qquad s = 1 \text{ or } 2$$

Now let's do the other situation where our final state was a negative energy state. Everything is the same, except we get a +i instead of a -i due to the signs in our original equation above, which I repeat here.

$$\psi_f{}^{\dagger\prime}.\, S_F\, e\not{A}\, \Psi \quad = \psi_f{}^{\dagger\prime}.\{ -i\theta \int d^3p\, \Sigma_{r=1,2} \psi_p{}^{r\prime}\; \bar{\psi}_p{}^r + i\theta_b \int d^3p\, \Sigma_{r=3,4} \psi_p{}^{r\prime}\; \bar{\psi}_p{}^r \}\, e\not{A}\, \Psi$$

So here are the two results right next to each other: ( could use $\varepsilon_f$ to write as one equation)

$$S_{fi} \;=\; \delta_{fi} \;- i\, \bar{\psi}_f{}^s\, e\not{A}\, \Psi \qquad s=1,2 \quad \text{form 5} \qquad 6.56$$

$$S_{fi} \;=\; \delta_{fi} \;+ i\, \bar{\psi}_f{}^s\, e\not{A}\, \Psi \qquad s=3,4 \quad \text{form 5} \qquad 6.56$$

where

$$\bar{\psi}_f{}^s \;= (\sqrt{m/E})\, (1/2\pi)^{3/2}\, e^{+ipf.x}\; \bar{w}^s(p_f)$$

Now we can obtain our "final two forms" for $S_{fi}$. The first requires (0) which is : $\theta\, \Psi' = \pm\, iS_F'\, \gamma^0.\, \psi$ where the ± depends on which way you are going in time. I think we can say $\theta\, \Psi' = \varepsilon_i\, iS_F'\, \gamma^0.\, \psi$ if this represents motion of our initial state, then we get

$S_{fi} = \delta_{fi} - \varepsilon_s\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \Psi$ form 6 6.33a f5 + (0)

$= \delta_{fi} - \varepsilon_s\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \{\varepsilon_i\, iS_F'\, \gamma^{\circ}.\, \psi\} = \delta_{fi} + \varepsilon_f\varepsilon_i \bar{\psi}_f^{s}\, e\not{A} S_F'\, \gamma^{\circ}.\, \psi_i$ form 6

Now finally let's go for form 7 which starts again with form 5 [$\Psi' = \psi' + S_F\, e\not{A}\, \Psi$ is (3) ]

$S_{fi} = \delta_{fi} - \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \Psi$ form 7 f5 + (3) twice

$= \delta_{fi} - \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \{\psi' + S_F\, e\not{A}\, [\psi' + S_F\, e\not{A}\, \Psi]\}$

$= \delta_{fi} - \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \psi_i - \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, S_F\, e\not{A}\, \psi_i + \ldots$ form 7

Now let's gather it all up:

| | | |
|---|---|---|
| $\theta\, \Psi' = \varepsilon\, iS_F'\, \gamma^{\circ}.\, \psi$ | (0) | forward in time $\theta(t'-t)$ and $\varepsilon = 1$, else both opposite |
| $\theta\, \psi' = \varepsilon\, iS_F\, \gamma^{\circ}.\psi$ | (1) | **6.49** and **6.50** |
| $S_F' = S_F + S_F\, e\not{A}\, S_F'$ | (2) | **6.51** |
| $\Psi' = \psi' + S_F\, e\not{A}\, \Psi$ | (3) | **6.53** |

$S_{fi} = \psi_f^{\dagger}.\, \Psi_i^{(+)}$ form 1 defines $S_{fi}$

$S_{fi} = i\, \psi_f^{\dagger}.S_F'\, \gamma^{\circ}.\, \psi_i$ form 2 f1 + (0)

$S_{fi} = \delta_{fi} + i\, \psi_f^{\dagger}.\, S_F\, e\not{A}\, S_F'\gamma^{\circ}.\, \psi_i$ form 3 f2 + (2) + (1)

$S_{fi} = \delta_{fi} + \psi_f^{\dagger}.\, S_F\, e\not{A}\, \Psi$ form 4 f1 + (3)

$S_{fi} = \delta_{fi} + \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \Psi$ form 5 **6.56** special derivation

$S_{fi} = \delta_{fi} + \varepsilon_f\varepsilon_i \bar{\psi}_f^{s}\, e\not{A} S_F'\, \gamma^{\circ}.\, \psi_i$ form 6 f5 + (0)

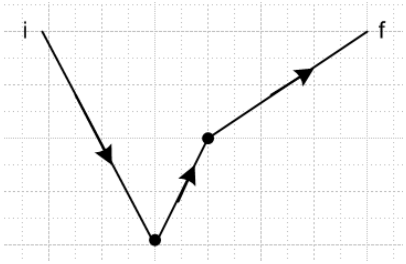
$S_{fi} = \delta_{fi} - \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, \psi_i - \varepsilon_f\, i\, \bar{\psi}_f^{s}\, e\not{A}\, S_F\, e\not{A}\, \psi_i + \ldots$ form 7 **6.57** $\varepsilon_f = 1$ f5 + (3) twice

Most of the above "forms" don't even appear in Chapter 6, mainly those involving $S_F'$, the full propagator. End of Very Long Note. Resuming original note flow after the following bar.

---

Now comes the clincher: by selecting your i and f states "properly", you find that you can be talking about four different processes! They are all "scattering" off the potential one or more times, the only question is how you interpret this scattering.

Consider the following graph:




This is order n = 2 because there are two vertices. The "initial" state (upper left) is an electron with negative energy ($\varepsilon_i = -1$) and we indicate its quantum numbers as $-p_i$, $-s_i$ let's say, So we put that in for $\psi_i$ in 6.57. Since this ne state is going backwards in time, we label it $\psi_i^{(-)}$ as in page 97 equation A. We know that in hole theory, this is really a positron going forward in time with $+p_i$ and $+s_i$. We saw how the two 4-vectors p and s change sign as you do the CPT on ψ. I would therefore be inclined to indicate the $\psi_i$ negative energy spinor by $v(-p_+, -s_+)$ and I would then interpret this branch as a positron going forward in time with $+ p_+$ and $+ s_+$. But they use opposite signs in equation A, I don't know why. I do agree with the long red paren sentence which matches my CPT comment just made, so you would think they would want $v(-p_+, -s_+)$, but that is not what they have written.

In the above graph for the final state $\psi_f$ we would put $u(p_-, s_-)$ and there is less confusion.

One possible explanation of our confusion regarding $\psi_i$ is this: (1) perhaps the sign change in $p_+$ is being compensated by the $\varepsilon_r = -1$ ; (2) it is true that v is defined with the spin reversed as shown in 3.16 on page 32. It does seem reasonable that they would set up the formalism so that v(p,s) would represent a positron with p and s.

Note Added: Here is how I would treat the above diagram. I assume the arrow indicates an electron flow. So $\psi_i$ is a neg energy state with $\varepsilon_i = -1$ and $\psi_f$ is a positive energy state with $\varepsilon_f = +1$. From form 7 above I would say you compute this diagram as follows:

$$S_{fi} = + \varepsilon_f \, i \, \bar{\psi}_f^{r'} \, e\not{A} \, S_F \, e\not{A} \, \psi^r_i$$

$$\bar{\psi}_f^{r'} = (\sqrt{m/E})\,(1/2\pi)^{3/2}\, e^{+i\,\varepsilon_f\, p_f .x}\; \bar{w}^r(p_f) \qquad r' = 2 \text{ or } 2 \qquad \varepsilon_f = +1$$

$$\psi_i^{r} = (\sqrt{m/E})\,(1/2\pi)^{3/2}\, e^{-i\,\varepsilon_i\, p_i .x}\; w^r(p_i) \qquad r = 3 \text{ or } 4 \qquad \varepsilon_i = -1$$

Note on the basic spinors.

Suppose we did this:

$$u(p,s) = B(p)\, R(\theta,\varphi)\, w^1(0) = \exp(-ip\,\hat{p}\bullet\mathbf{K})\exp(-i\hat{\mathbf{s}}\bullet\mathbf{J})\, w^1(0) = \exp(-i\mathbf{p}\bullet\mathbf{K})\, \exp(-i\mathbf{s}\bullet\mathbf{J})\, w^1(\mathbf{0})$$

where we are using 4x4 matrices from the Dirac representation. Then of course u(p,s) is a function of the two items as four-vectors. Then we would have the following four results (notation of 3.16 page 32).

$$u(p,s) = \exp(-i\mathbf{p}\bullet\mathbf{K})\, \exp(-i\mathbf{s}\bullet\mathbf{J})\, w^1(\mathbf{0})$$

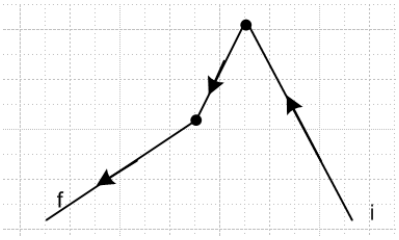
$u(p,-s) = \exp(-i\mathbf{p}\bullet\mathbf{K}) \exp(-i\mathbf{s}\bullet\mathbf{J})\, w^{2}(\mathbf{0})$

$v(p,s) = \exp(-i\mathbf{p}\bullet\mathbf{K}) \exp(-i\mathbf{s}\bullet\mathbf{J})\, w^{4}(\mathbf{0})$

$v(p,-s) = \exp(-i\mathbf{p}\bullet\mathbf{K}) \exp(-i\mathbf{s}\bullet\mathbf{J})\, w^{3}(\mathbf{0})$

I need to get this stuff cleaned up for the Dirac 4x4 representation, it goes on the list. Other books I notice just use the basic spinors above.

Apart from this Feynman rule detail, we can view the above graph as a scattering graph rotated 90 degrees. In similar fashion, we have "pair annihilation" represented by this graph




Where "i" is a normal u(p,s) incoming electron, and "f" is a negative energy electron v(p,s).

The two other "rotations" give us "electron scattering" and "positron scattering" .

Comments:

(1) There are no photons anywhere, just $\mathbb{A}$. The EM field is not quantized at this point in our progress through both BD books. The vertex gets a classical $A^{\mu}$ potential, although it has been "installed into the theory" in a relativistically correct manner. HOWEVER: photons and photon propagators are going to suddenly appear in the upcoming chapter 7 so we can compute all the basic lowest order processes.

(2) It seems that the propagators can carry information at an arbitrary velocity, even greater than c. I say this because the theta functions are of the form $\theta(t-t')$ with no account made for flight time.

(3) A propagator Green's function can be represented as a sum over all paths of $e^{iS}$ where S is the action. This whole concept, also invented by Feynman it seems (1948), has not been mentioned by BD up to this point. I think this $e^{iS}$ method is "the modern method" of doing field theory now, but I am reading a 1965 book. Goldstein (1950) mentioned the action S as the time integral of L, but he never got it up into an exponent. At some point I will want to get into this subject, but don't want to go off on yet another tangent right now.