1 Lorentz Transformations and Lorentz Groups

1.1 Lorentz transformations

1.1.1 Introduction

Einstein's starting point: *the speed of light in all inertial systems is constant.* Consider the two coordinate systems **O** and **O′** with **O′** moving along the $+x$ axis with the speed $\underline{v}$:

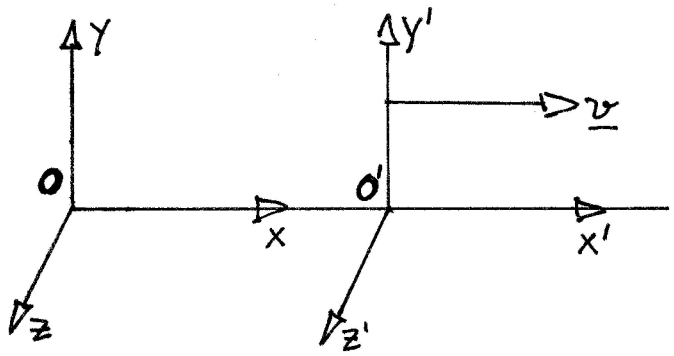

Figure 1.1: The coordinate systems **O** and **O′**

Let $x = x'$ for $t = t' = 0$. Then

$$\begin{cases} y' = y \\ z' = z \\ x' = \dfrac{x - vt}{\sqrt{1 - v^2/c^2}} \\ t' = \dfrac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \end{cases} \tag{1.1}$$

The inverse transformation is obtained by changing the sign of v:

$$\begin{cases} x = \dfrac{x' + vt}{\sqrt{1 - v^2/c^2}} \\ t = \dfrac{t' + vx'/c^2}{\sqrt{1 - v^2/c^2}} \end{cases} \tag{1.2}$$

Assuming that a light signal from **O** obeys

$$x^2 + y^2 + z^2 = c^2 t^2 \tag{1.3}$$

we find indeed that

$$x'^2 + y'^2 + z'^2 = c^2 t'^2 \tag{1.4}$$

The speed of light is the same in **O** and **O′**.

— Recall the **Michelson-Morley** experiment (1887).

— Eq. (1.1) was first introduced by **H. A. Lorentz** (1892, 1895). Its use in relativity is due to Einstein.

— Recall **Fitzgerald** (1895): lengths change! (Lorentz: mass changes).

The distance between points 1 and 2:

$$\begin{cases} \Delta x_i = x_i^{(1)} - x_i^{(2)}, \quad i = 1, 2, 3 \\ \Delta t = t^{(1)} - t^{(2)} \end{cases} \tag{1.5}$$

In the coordinate system **O**:

$$\sum_{i=1}^{3} (\Delta x_i)^2 - c^2 (\Delta t)^2 = 0 \tag{1.6a}$$

In the coordinate system $\mathbf{O}'$:

$$\sum_{i=1}^{3} (\Delta x_i')^2 - c^2 (\Delta t')^2 = 0 \tag{1.6b}$$

A coordinate transformation $\mathbf{O} \leftrightarrow \mathbf{O}'$ obeying (1.6) is called a **Lorentz transformation**.

Defining

$$l = ct$$

$$x_4 = \mathrm{i}l = \mathrm{i}ct$$

we can write

$$\Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2 - \Delta l^2 = 0 \tag{1.6c}$$

$$\Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2 + \Delta x_4^2 = 0 \tag{1.6d}$$

More generally one can consider the transformation

$$\Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2 + \Delta x_4^2 = s^2 \tag{1.7}$$

The transformations with $s \neq 0$ lie outside the *"light cone"*.

The general linear transformation becomes[1]

$$x_\mu' = a_\mu + b_{\mu\alpha} x_\alpha \tag{1.8}$$

[1]From now on Greek indices, $\alpha, \beta, \ldots \mu, \nu$ means the indices run from 1–4, while Latin indices $i, j, \ldots$ run from 1–3. If the same index (for example α) appears twice a summation ($\sum_\alpha$) is to be understood.

1.1.2 Rotations and translations of the four-dimensional space

Theorem : For pure rotations, $a_\mu = 0$, the transformation (1.8) will satisfy (1.7) iff (*if and only if*)[2]

$$b_{\mu\alpha}b_{\nu\alpha} = b_{\alpha\mu}b_{\alpha\nu} = \delta_{\mu\nu} \tag{1.9}$$

Proof: For the the light cone, $s^2 = 0$, (1.6) $\Rightarrow$

$$\left.\begin{aligned} \sum_\mu (x_\mu^{(1)})^2 &= \sum_\mu (x_\mu'^{(1)})^2 \\ \sum_\mu \Delta x_\mu^2 = \sum_\mu (x_\mu^{(1)} - x_\mu^{(2)})^2 &= \sum_\mu (x_\mu'^{(1)} - x_\mu'^{(2)})^2 \end{aligned}\right\} \Rightarrow$$

$$\sum_\mu x_\mu^{(1)} x_\mu^{(2)} = \sum_\mu x_\mu'^{(1)} x_\mu'^{(2)} \tag{1.10}$$

Inserting (1.8) on the RHS (*right hand side*), we get

$$\begin{aligned} \sum_\mu (\sum_\alpha b_{\mu\alpha} x_\alpha^{(1)})(\sum_\beta b_{\mu\beta} x_\beta^{(2)}) &= \sum_{\alpha\beta} (\sum_\mu b_{\mu\alpha} b_{\mu\beta}) x_\alpha^{(1)} x_\beta^{(2)} \\ &\overset{\text{LHS}}{=} \sum_\alpha x_\alpha^{(1)} x_\alpha^{(2)} \end{aligned}$$

This must hold for all $x_\alpha^{(1)}, x_\beta^{(2)} \Rightarrow$

$$\sum_\mu b_{\mu\alpha} b_{\mu\beta} = \delta_{\alpha\beta} \qquad \square$$

As x_1–x_3 are real and x_4 is pure imaginary, all other a_μ and $b_{\mu\alpha}$ are real, except $a_4, b_{41}, b_{42}, b_{43}, b_{14}, b_{24}$ and b_{34} which are pure imaginary.

1.1.3 Lorentz transformations as rotations

For a rotation through ϕ around x_3,

[2]The δ denotes **Kronecker's** delta

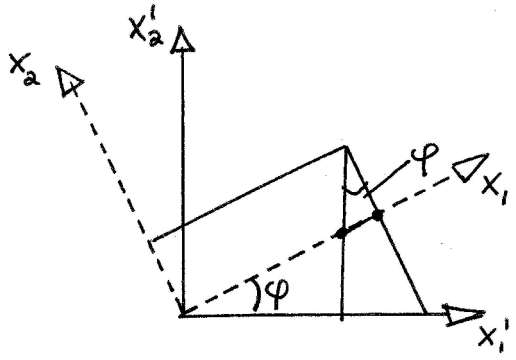

Figure 1.2: Rotation through ϕ

$$\begin{cases} x_1' = x_1 \cos\phi - x_2 \sin\phi \\ x_2' = x_1 \sin\phi + x_2 \cos\phi \\ x_3' = x_3 \\ x_4' = x_4 \end{cases} \tag{1.11}$$

Similarly, the Lorentz transformation (1.1) becomes

$$\begin{cases} x_1' = x_1 \cos\psi - x_4 \sin\psi \\ x_2' = x_2 \\ x_3' = x_3 \\ x_4' = x_1 \sin\psi + x_4 \cos\psi \end{cases} \tag{1.12}$$

Now ψ is imaginary, $\sin\psi$ is imaginary and $\cos\psi$ is real. (For real ψ, set $l = ct = -\mathrm{i}x_4, \quad x_4 = \mathrm{i}l$)

$$\begin{cases} x_1' = x_1 \cos\psi - \mathrm{i}l \sin\psi \\ l' = -\mathrm{i}x_1 \sin\psi + l \cos\psi \end{cases} \tag{1.13}$$

For the origin $\mathbf{O'}(\mathbf{x_1'} = \ldots = \mathbf{l'} = \mathbf{0})$ we get

$$x_1 = vt = v\frac{l}{c} = \frac{v}{c}l \equiv \beta l$$

Hence (1.13a) $\Rightarrow$

$$\begin{aligned} x_1 = \mathrm{i}l\frac{\sin\psi}{\cos\psi} = \beta l \Rightarrow \\ \beta = \mathrm{i}\tan\psi \end{aligned} \tag{1.14}$$

$$\begin{cases} \sin\psi = \dfrac{1}{\sqrt{1+\cot^2\psi}} = \dfrac{1}{\sqrt{1-\beta^{-2}}} = \dfrac{-\mathrm{i}\beta}{\sqrt{1-\beta^2}} \\ \cos\psi = \dfrac{1}{\sqrt{1+\tan^2\psi}} = \dfrac{1}{\sqrt{1-\beta^2}} \end{cases} \tag{1.15}$$

Inserting (1.15) in (1.13), we get

$$\begin{cases} x_1' = \dfrac{x_1 - \mathrm{i}l(-\mathrm{i}\beta)}{\sqrt{1-\beta^2}} = \dfrac{x_1 - vt}{\sqrt{1-\beta^2}} \\ l' = ct' = \dfrac{-\mathrm{i}x_1(-\mathrm{i}\beta) + l}{\sqrt{1-\beta^2}} = \dfrac{l - \beta x_1}{\sqrt{1-\beta^2}} \Rightarrow \\ t' = \dfrac{t - x_1 v/c^2}{\sqrt{1-\beta^2}} \end{cases} \tag{1.16}$$

in agreement with (1.1). *The invariance of c yields* (1.1) *!*

1.1.4 Addition of velocities

For two subsequent Lorentz transformations, corresponding to velocitites $\underline{v_1}$ and $\underline{v_2}$, we get

$$\begin{aligned} \beta_{12} &= \mathrm{i}\tan(\psi_1 + \psi_2) = \mathrm{i}\frac{\tan\psi_1 + \tan\psi_2}{1 + \tan\psi_1 \tan\psi_2} \\ &= \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2} \end{aligned} \tag{1.17}$$

Here $\beta_i = v_i/c$. For $\underline{u} = \underline{u}' + \underline{v}$, we have

$$u = \frac{u' + v}{1 + u'\frac{v}{/}c^2} \tag{1.18}$$

cf. eq. (1.9) in Pyykkö (1975). Classically we would simply have $u = u' + v$, but with (1.18) we see that for $u' \to c, u \to \frac{c+v}{1+v/c} = c$. Even if we would have $v \to c$ at the same time, we get $u \to \frac{c+c}{1+c/c} = c$. The speed of light is not exceeded.

1.1.5 Perpendicular motion

Let now $\mathbf{O}'$ move in the x-direction and a body move in the y-direction:

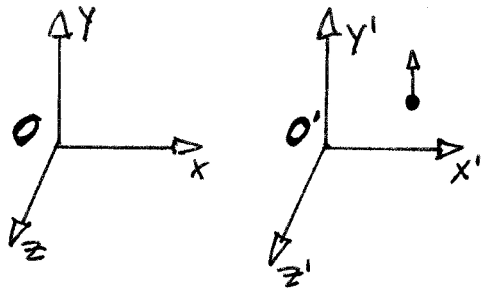

Figure 1.3: **O** and $\mathbf{O}'$, again

Consider two different times, t'_1 and t'_2. Then the velocity in the $\mathbf{O}'$ frame becomes

$$u'_y = \frac{\Delta y'}{\Delta t'} = \frac{y'_2 - y'_1}{t'_2 - t'_1} \tag{1.19}$$

With eq. (1.1) we get

$$\begin{cases} y'_2 - y'_1 = y_2 - y_1 \\ t'_2 - t'_1 = \dfrac{t_2 - t_1 - (x_2 - x_1)v/c^2}{\sqrt{1-\beta^2}} = \dfrac{\Delta t - \Delta x(v/c^2)}{\sqrt{1-\beta^2}} \end{cases} \tag{1.20}$$

$$\frac{\Delta y'}{\Delta t'} = \frac{\Delta y\sqrt{1-\beta^2}}{\Delta t - \Delta x(v/c^2)} = \frac{\Delta y}{\Delta t}\frac{\sqrt{1-\beta^2}}{1 - \frac{\Delta x}{\Delta t}v/c^2} \tag{1.21}$$

$$u'_y = \frac{u_y\sqrt{1-\beta^2}}{1 - u_x(v/c^2)} \tag{1.22}$$

By changing the sign of v and primed/unprimed quantities, we can write

$$u_y = \frac{u'_y\sqrt{1-\beta^2}}{1 + u'_x(v/c^2)} \tag{1.23}$$

1.1.6 The relativistic mass transformation

Consider an elastic collision between two bodies having the mass m. Let frame $\mathbf{O}$ have $u_{xA} = 0$ and let $\mathbf{O}'$ be the centre-of-mass frame.

Small letters before impact,
Capital letters after impact

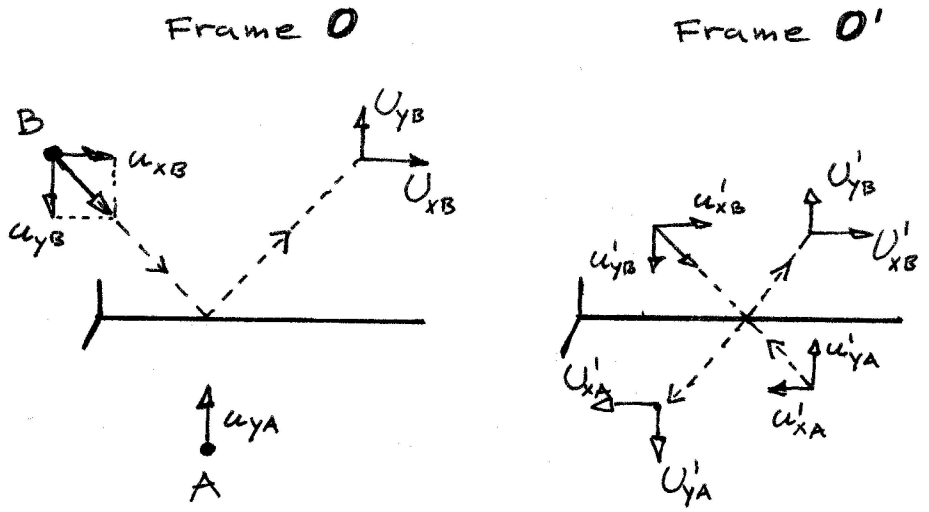

Figure 1.4: $v = u'_{xB} = -u'_{xA}$

From (1.22), in the c.m. coordinates,

$$\begin{cases} u'_{yB} = \dfrac{u_{yB}\sqrt{1-\beta^2}}{1-u_{xB}v/c^2} \\ u'_{yA} = u_{yA}\sqrt{1-\beta^2} \qquad (\text{as } u_{xA}=0) \end{cases} \tag{1.24}$$

In the c.m. system $|y'_{yA}| = |y'_{yB}| \Rightarrow$

$$u_{yA} = \frac{u_{yB}}{1-u_{xB}v/c^2} \tag{1.25}$$

The laws of nature have to be the same in all inertial frames. Putting the relativistic momentum changes equal in the **O** frame,

$$2m_A u_{yA} = 2m_B u_{yB} \tag{1.26}$$

we get

$$m_B = m_A \frac{u_{yA}}{u_{yB}} = \frac{m_A}{1-u_{xB}v/c^2} \tag{1.27}$$

Note that the two masses are allowed to be different. As

$$v = u'_{xB} = \frac{u_{xB}-v}{1-u_{xB}v/c^2} \Rightarrow \tag{1.28}$$

$v^2 - \frac{2c^2}{u_{xB}}v + v^2 = 0$,

$$v = \frac{c^2}{u_{xB}} \overset{(+)}{-} \sqrt{(\frac{c^2}{u_{xB}})^2 - c^2} \overset{v<c}{=} \frac{c^2}{u_{xB}}\left[1 - \sqrt{1-(\frac{u_{xB}}{c})^2}\right]. \tag{1.29}$$

Inserting this in (1.27), and letting $u_{yA}, u_{yB} \to 0$,

$$\begin{aligned} m_B &= \frac{m_A}{1-\frac{u_{xB}}{c^2}\frac{c^2}{u_xB}\left[1-\sqrt{1-(\frac{u_{xB}}{c})^2}\right]} = \frac{m_A}{\sqrt{1-(\frac{u_{xB}}{c})^2}} \\ &= \frac{m_A}{\sqrt{1-\beta^2}} \end{aligned} \tag{1.30}$$

In three dimensions,

$$\underline{p} = m_0 \frac{\underline{u}}{\sqrt{1-u^2/c^2}} \tag{1.31}$$

According to Einstein, *The Meaning of Relativity*, p. 45, H. A. Lorentz used $\underline{F} = \mathrm{d}\underline{p}/\mathrm{d}t$ with this $\underline{p}$. $\underline{F} \neq m\underline{a}$.

1.1.7 Derivation of $E = mc^2$

The kinetic energy, T, is the work done by a force, F, in accelerating a particle to the speed u:

$$\begin{aligned} T &= \int_{u=0}^{u=u} F\mathrm{d}x = \int_{u=0}^{u=u} \frac{\mathrm{d}}{\mathrm{d}t}(mu)\mathrm{d}x = \int_{u=0}^{u=u} \mathrm{d}(mu)\frac{\mathrm{d}x}{\mathrm{d}t} \\ &= \int_{u=0}^{u=u} (m\mathrm{d}u + u\mathrm{d}m)u = \int_{u=0}^{u=u} (mu\mathrm{d}u + u^2\mathrm{d}m) \end{aligned} \tag{1.32}$$

Recalling that

$$\begin{aligned} m = \frac{m_0}{\sqrt{1-u^2/c^2}} \Rightarrow \\ m^2c^2 - m^2u^2 = m_0^2c^2 \end{aligned}$$

and differentiating

$$\begin{aligned} 2mc^2\mathrm{d}m - m^2 \cdot 2u\mathrm{d}u - u^2 \cdot 2m\mathrm{d}m = 0 \quad \Big| : 2m \\ c^2\mathrm{d}m = mu\mathrm{d}u + u^2\mathrm{d}m \end{aligned} \tag{1.33}$$

which is the integrand (1.32)! $\Rightarrow$

$$T = \int_{u=0}^{u=u} c^2\mathrm{d}m = c^2 \int_{m=m_0}^{m=m} \mathrm{d}m = mc^2 - m_0c^2 \tag{1.34}$$

Equivalently,

$$T = m_0c^2\left[\frac{1}{\sqrt{1-u^2/c^2}} - 1\right] \tag{1.35}$$

Denoting the total energy mc^2 by E,

$$E = m_0c^2 + T \tag{1.36}$$

where m_0c^2 is the *rest energy*

— **Classical mechanics:** The energy is defined apart from a constant.

— **Relativistic mechanics:** This constant is m_0c^2.

1.1.8 Connection between T and p

From (1.35),

$$\begin{aligned} (T + m_0c^2)^2 \quad &= \tfrac{(m_0c^2)^2}{1-u^2/c^2} \quad \Rightarrow \\ 1 - \frac{u^2}{c^2} \quad &= \tfrac{(m_0c^2)^2}{(T+m_oc^2)^2} \Rightarrow \end{aligned} \tag{1.37}$$

$$u = c\sqrt{1 - \frac{(m_0c^2)^2}{(T + m_0c^2)^2}} \tag{1.38}$$

$\Rightarrow$

$$\begin{aligned} p &= \tfrac{m_0 u}{\sqrt{1-u^2/c^2}} = \tfrac{m_0}{\sqrt{1-u^2/c^2}} c\sqrt{1 - \tfrac{(m_0c^2)^2}{(T+m_0c^2)^2}} \\ &\overset{(1.37)}{=} m_0 c\sqrt{1 - \tfrac{(m_0c^2)^2}{(T+m_0c^2)^2}} \times \tfrac{T+m_0c^2}{m_0c^2} \\ &= \tfrac{1}{c}\sqrt{(T + m_o c^2)^2 - (m_0c^2)^2} \Rightarrow \end{aligned} \tag{1.39}$$

$$\begin{aligned} p^2c^2 &= (T + m_0c^2)^2 - (m_0c^2)^2 \Rightarrow \\ E^2 &= (T + mo_c^2)^2 = p^2c^2 + m_0^2c^4 \end{aligned} \tag{1.40}$$

1.2 Lorentz matrices

Consider the vector

$$\underline{x_\mu} = (x_1, x_2, x_3, x_4) \tag{1.41}$$

with $x_1 = x, x_2 = y, x_3 = z, x_4 = \mathrm{i}x_0 = \mathrm{i}ct$. Introducing

$$\beta = \frac{v}{c},\ \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{1.42}$$

we can write the transformation (1.16) as

$$\begin{cases} x_1' = \dfrac{x - vt}{\sqrt{1-\beta^2}} & = \gamma(x_1 + \mathrm{i}\beta x_4) \\ x_2' = x_2 & \\ x_3' = x_3 & \\ x_4' = \mathrm{i}ct' = \mathrm{i}c\dfrac{t - xv/c^2}{\sqrt{1-\beta^2}} & = \gamma(x_4 - \mathrm{i}\beta x_1) \end{cases} \tag{1.43}$$

Using the Einstein summation notation,

$$x_\mu' = a_{\mu\nu}x_\nu \tag{1.44}$$

we have

$$a_{\mu\nu} = \begin{pmatrix} \gamma & 0 & 0 & \mathrm{i}\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{i}\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \tag{1.45}$$

We observe that

a. The length of the vector remains invariant:

$$\begin{aligned} x'_\mu x'_\mu &= \gamma^2(x_1 + \mathrm{i}\beta x_4)^2 + x_2^2 + x_3^2 + \gamma^2(x_4 - \mathrm{i}\beta x_1)^2 \\ &= x_1^2 \underbrace{\gamma^2(1-\beta^2)}_{1} + x^2 + x_3^2 + x_4^2\gamma^2(1-\beta^2) \\ &= x_\mu x_\mu \end{aligned} \tag{1.46}$$

b. The reality properties of the components are preserved: x_1, x_2, x_3 are real and x_4 is imaginary.

Definition: $a_{\mu\nu}$ is a **homogenous Lorentz transformation** if

$$x'_\mu = a_{\mu\nu} x_\nu \tag{1.47}$$

leaves $x_\mu x_\mu$ invariant and preserves the reality properties of x_μ.

We saw on page 3 that $a_{\mu\nu}$ is orthogonal,

$$\begin{aligned} a_{\nu\mu} a_{\lambda\mu} &= \delta_{\nu\lambda} \\ a_{\mu\nu} a_{\mu\lambda} &= \delta_{\nu\lambda} \end{aligned} \tag{1.48}$$

its rows and columns are orthogonal.

Corollary:

$$\det a = \pm 1 \tag{1.49}$$

Example 1: Rotations in the 3-D space are Lorentz transformations:

$$x'_k = a_{kl} x_l, \quad x'_4 = x_4 \tag{1.50}$$

Here $x_k x_k$ remains invariant (rotation!).

Definition: Lorentz transformations with $a_{44} > 0$ are called *orthochronous*. Orthochronous Lorentz transformations with

$$\det a = +1 \tag{1.51}$$

are called *proper* Lorentz transformations

Example 2: For the particular case (1.45),

$$\det a = \det \begin{vmatrix} \gamma & 0 & 0 & \mathrm{i}\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{i}\gamma\beta & 0 & 0 & \gamma \end{vmatrix}$$
$$= \gamma \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \gamma \end{vmatrix} - \mathrm{i}\gamma\beta \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\mathrm{i}\gamma\beta & 0 & 0 \end{vmatrix} = \gamma^2 + (\mathrm{i}\gamma\beta)^2$$
$$= \gamma^2(1-\beta^2) = 1 \tag{1.52}$$

$a_{44} = \gamma > 1 \Rightarrow a$ is a proper LT.

Example 3: The *inversion*

$$a_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \tag{1.53}$$

is an orthochronous but *improper* LT: $a_{44} > 0, \quad \det a = -1$.

1.3 Infinitesimal Lorentz transformations

Consider the transformations

$$x'_\mu = a_{\mu\nu} x_\nu, \quad a_{\mu\nu} = \delta_{\mu\nu} + \epsilon_{\mu\nu} \tag{1.54}$$

where the $\epsilon_{\mu\nu}$ are infinitesimal quantities. The orthogonality requirement (1.48) $\Rightarrow$

$$(\delta_{\mu\nu} + \epsilon_{\mu\nu})(\delta_{\mu\lambda} + \epsilon_{\mu\lambda}) = \delta_{\nu\lambda} \tag{1.55}$$

Writing out the above multiplication, we get:

$$\underbrace{\delta_{\mu\nu}\delta_{\mu\lambda}}_{\delta_{\nu\lambda}} + \underbrace{\delta_{\mu\nu}\epsilon_{\mu\lambda}}_{\epsilon_{\nu\lambda}} + \underbrace{\epsilon_{\mu\nu}\delta_{\mu\lambda}}_{\epsilon_{\lambda\nu}} + \underbrace{\epsilon_{\mu\nu}\epsilon_{\mu\lambda}}_{\approx 0} = \delta_{\nu\lambda}$$

$$\Rightarrow \epsilon_{\nu\lambda} = -\epsilon_{\lambda\nu} \tag{1.56}$$

where we have assumed the product of two infinitesimals to be negligible.

The reality properties demand that

$$\epsilon^*_{ik} = \epsilon_{ik}, \quad \epsilon^*_{44} = \epsilon_{44}, \quad \epsilon^*_{i4} = -\epsilon_{i4}, \quad \epsilon^2_{4k} = -\epsilon_{4k} \tag{1.57}$$

Hence the ϵ matrix becomes

$$\epsilon_{\mu\nu} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 & \mathrm{i}\lambda_1 \\ -\omega_3 & 0 & \omega_1 & \mathrm{i}\lambda_2 \\ \omega_2 & -\omega_1 & 0 & \mathrm{i}\lambda_3 \\ -\mathrm{i}\lambda_1 & -\mathrm{i}\lambda_2 & -\mathrm{i}\lambda_3 & 0 \end{pmatrix} \tag{1.58}$$

The vectors

$$\begin{aligned} \underline{\omega} &= (\omega_1, \omega_2, \omega_3) \\ \underline{\lambda} &= (\lambda_1, \lambda_2, \lambda_3) \end{aligned} \tag{1.59}$$

are real and infinitesimal but otherwise arbitrary.

Example 4: Let $\underline{\omega} = (\omega_1, \omega_2, \omega_3), \quad \underline{\lambda} = (0, 0, 0) \Rightarrow$

$$a_{\mu\nu} = \begin{pmatrix} 1 & \omega_3 & -\omega_2 & 0 \\ -\omega_3 & 1 & \omega_1 & 0 \\ \omega_2 & -\omega_1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.60}$$

This is a three-dimensional (3-D) rotation where $\omega_1, \omega_2, \omega_3$ are the rotation angles about the three axes.

Example 5: Let $\underline{\omega} = (0, 0, 0), \quad \underline{\lambda} = (\lambda_1, 0, 0) \Rightarrow$

$$a_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & \mathrm{i}\lambda_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{i}\lambda_1 & 0 & 0 & 1 \end{pmatrix} \tag{1.61}$$

Suppose that $v_1/c \ll 1$. Then, recalling (1.45), for

$$a_{14} = \mathrm{i}\gamma\beta$$

$$\mathrm{i}\gamma\beta = \mathrm{i}\frac{v/c}{\sqrt{1 - v^2/c^2}} \approx \mathrm{i}\frac{v_1}{c} + \mathbf{O}((\frac{v_1}{c})^3) \tag{1.62}$$

Thus (1.61) is of the form (1.45), we have a translational Lorentz transformation in the x-direction.

The *generators* of infinitesimal Lorentz transformations are the six ma-

trices

$$\Omega_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Omega_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Omega_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.63}$$

$$\Lambda_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \Lambda_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{1.64}$$

Then we can express

$$\epsilon_{\mu\nu} = \underline{\omega} \cdot \underline{\Omega} + \mathrm{i}\underline{\lambda} \cdot \underline{\Lambda} \tag{1.65}$$

Theorem : The commutators

$$[\Omega_i, \Omega_k] \equiv \Omega_i \Omega_k - \Omega_k \Omega_i, \quad \text{etc.} \tag{1.66}$$

obey

$$\begin{aligned} [\Omega_i, \Omega_k] &= -\delta_{ikl}\Omega_l \\ [\Lambda_i, \Lambda_k] &= -\delta_{ikl}\Omega_l \\ [\Lambda_i, \Omega_k] &= -\delta_{ikl}\Lambda_l \end{aligned} \tag{1.67}$$

where the *permutation symbol*

$$\delta_{ikl} = \begin{cases} +1, & \text{if } i,k,l \text{ all different, } (ikl) \text{ even permutation} \\ -1, & \text{if } i,k,l \text{ all different, } (ikl) \text{ odd permutation} \\ 0, & \text{otherwise} \end{cases} \tag{1.68}$$

Proof: Explicit calculation. E.g.

$$\Omega_1\Omega_2 = \begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}, \Omega_2\Omega_1 = \begin{pmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix},$$

$$[\Omega_1, \Omega_2] = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = -\Omega_3, \quad (123) \text{ even} \qquad \square$$

In a more concise form we can express (1.67) for the six generators

$$I_s = (\Omega_1, \Omega_2, \Omega_3, \Lambda_1, \Lambda_2, \Lambda_3) \tag{1.69}$$

as

$$[I_r, I_s] = C^t_{rs} I_t \tag{1.70}$$

where the C^t_{rs} are called *structure constants*.

1.4 The Lorentz group

The six generators $\Omega_1, \Omega_2, \Omega_3, \Lambda_1, \Lambda_2, \Lambda_3$ generate the **proper Lorentz group**, a *continuous group*, whose elements are

$$\beta_s = (\omega_1, \omega_2, \omega_3, \mathrm{i}\lambda_1, \mathrm{i}\lambda_2, \mathrm{i}\lambda_3) \tag{1.71}$$

The 3D rotations (1.60) form a subgroup. Adding the inversion

$$a_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \tag{1.72}$$

we get the **orthocronous Lorentz group**. Adding to it the *time inversion*

$$a_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \tag{1.73}$$

we get the **full Lorentz group** (with eight generators in all).

2 Relativistic Fields

2.1 Definition

A wave function spanning a representation of the proper Lorentz group is called a *relativistic field*. The only thing required is that

$$[I_r, I_s] = C_{rs}^t I_t \qquad [1.70]$$

must hold

In other words, $\psi_\alpha(x)$ is a relativistic field if it transforms according to

$$\psi'_\alpha(x') = [1 + \underline{\omega} \cdot \underline{\Omega} + \mathrm{i}\underline{\lambda} \cdot \underline{\Lambda}]_{\alpha\beta}\psi_\beta(x) \qquad (2.1)$$

For pure (3D) rotations,

$$\psi'(x') = \underbrace{[1 + \underline{\omega} \cdot \underline{\Omega}]}_{n\times n \text{ matrix}} \underbrace{\psi(x)}_{n\text{D vector}} \qquad (2.2)$$

The value of n is so far open.

On the other hand, the wave function for a particle with spin $\underline{S}$ transforms under a rotation $\underline{\omega}$ of the coordinate axes as

$$\psi'(x') = [1 + \frac{\mathrm{i}}{\hbar}\underline{\omega} \cdot \underline{S}]\psi(x) \qquad (2.3)$$

Hence the spin angular momentum

$$\underline{S} = -\mathrm{i}\hbar\underline{\Omega} \qquad (2.4)$$

for whatever $\underline{\Omega}$ is chosen.

A wave equation transforming according to a given representation of the Lorentz group describes a particle with a given spin.

2.2 Scalar fields

For one-component wave functions, the $\underline{\Omega}$ and $\underline{\Lambda}$ are 1×1 matrices. Eq. (1.70) $\Rightarrow \Omega_1\Omega_2 - \Omega_2\Omega_1 = 0 \Rightarrow \Omega_3 = 0$, cycl. Thus

$$\underline{\Omega} = \underline{\Lambda} = 0 \qquad (2.5)$$

(2.1) $\Rightarrow$

$$\psi'(x') = \psi(x) \qquad (2.6)$$

(2.4) $\Rightarrow$

$$\underline{S} = 0 \qquad (2.7)$$

Thus a scalar (one-component) field describes a spin-zero particle.

2.3 An $\underline{S} = 1$ field

Consider a four-component wave function A, transforming as

$$A'_\mu(x') = [1 + \underline{\omega} \cdot \underline{\Omega} + \mathrm{i}\underline{\lambda} \cdot \underline{\Lambda}]_{\mu\nu} A_\nu(x) \tag{2.8}$$

with the 4×4 representation in (1.63). Eq. (2.4)$\Rightarrow$

$$\begin{aligned} \underline{S} &= -\hbar^2 \left[\begin{pmatrix} 0 & & & \\ & -1 & & \\ & & -1 & \\ & & & 0 \end{pmatrix} + \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & -1 & \\ & & & 0 \end{pmatrix} + \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \right] \\ &= 2\hbar^2 \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = 2\hbar^2 \underline{\underline{\mathbf{1}}} \end{aligned} \tag{2.9}$$

$$S(S+1) = 2 \Rightarrow S = 1 \tag{2.10}$$

$S_z = -1, 0, 1$. This is the *self-representation* of the proper Lorentz group.

2.4 Two-component spinor fields

Can we find six 2×2 matrices, satisfying (1.67)? Try the Pauli matrices $\underline{\sigma}$, which satisfy

$$[\sigma_i, \sigma_k] = 2\mathrm{i}\delta_{ikl}\sigma_l \tag{2.11}$$

An explicit representation is

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.12}$$

Choose

$$\underline{\Omega} = \frac{\mathrm{i}}{2}\underline{\sigma} \tag{2.13}$$

$$\underline{\Lambda} = \pm\frac{\mathrm{i}}{2}\underline{\sigma} \tag{2.14}$$

We can verify that (1.67) holds

$$\begin{cases} [\Omega_i, \Omega_k] = -\dfrac{1}{4}[\sigma_i, \sigma_k] = -\dfrac{\mathrm{i}}{2}\delta_{ikl}\sigma_l = -\delta_{ikl}\Omega_l \\ [\Lambda_i, \Lambda_k] = -\dfrac{1}{4}[\sigma_i, \sigma_k] = -\dfrac{\mathrm{i}}{2}\delta_{ikl}\sigma_l = -\delta_{ikl}\Omega_l \\ [\Lambda_i, \Omega_k] = \quad \mp[\sigma_i, \sigma_k] = \mp\dfrac{\mathrm{i}}{2}\delta_{ikl}\sigma_l = -\delta_{ikl}\Lambda_l, \quad \text{indeed.} \end{cases} \tag{2.15}$$

Eq. (2.4) $\Rightarrow$

$$\underline{S} = -\mathrm{i}\hbar\underline{\Omega} = \frac{\hbar}{2}\underline{\sigma} \tag{2.16}$$

$$\underline{S}^2 = \frac{\hbar^2}{4}\Big[\begin{pmatrix}0&1\\1&0\end{pmatrix}^2 + \begin{pmatrix}0&-\mathrm{i}\\ \mathrm{i}&0\end{pmatrix}^2 + \begin{pmatrix}1&0\\0&-1\end{pmatrix}^2\Big] = \frac{3}{4}\hbar^2\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

$$S(S+1) = \frac{3}{4} \Rightarrow S = \frac{1}{2} \tag{2.17}$$

Consider then the inversion

$$P = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \tag{1.70}$$

For the matrices ER1.62–(1.64)

$$[P, \Omega] = 0 \tag{2.18}$$
$$\{P, \Lambda\} \equiv P\Lambda + \Lambda P = 0 \tag{2.19}$$

Example 1:

$$\left.\begin{aligned} P\Omega_1 &= \begin{pmatrix}-1&&&\\&-1&&\\&&-1&\\&&&1\end{pmatrix}\begin{pmatrix}&&1&\\&&&\\-1&&&\\&&&\end{pmatrix} = \begin{pmatrix}&&-1&\\&&&\\1&&&\\&&&\end{pmatrix} \\ \Omega_1 P &= \begin{pmatrix}&&1&\\&&&\\-1&&&\\&&&\end{pmatrix}\begin{pmatrix}-1&&&\\&-1&&\\&&-1&\\&&&1\end{pmatrix} = \begin{pmatrix}&&-1&\\&&&\\1&&&\\&&&\end{pmatrix} \end{aligned}\right\} \Rightarrow [P, \Omega_1] = 0$$

$$\left.\begin{aligned} P\Lambda_1 &= \begin{pmatrix}-1&&&\\&-1&&\\&&-1&\\&&&1\end{pmatrix}\begin{pmatrix}&&&1\\&&&\\&&&\\-1&&&\end{pmatrix} = \begin{pmatrix}&&&-1\\&&&\\&&&\\-1&&&\end{pmatrix} \\ \Lambda_1 P &= \begin{pmatrix}&&&1\\&&&\\&&&\\-1&&&\end{pmatrix}\begin{pmatrix}-1&&&\\&-1&&\\&&-1&\\&&&1\end{pmatrix} = \begin{pmatrix}&&&1\\&&&\\&&&\\1&&&\end{pmatrix} \end{aligned}\right\} \Rightarrow \{P, \Lambda_1\} = 0 \quad \square$$

Now , for the two-component case, either $\underline{\Omega} = \underline{\Lambda}$ or $\underline{\Omega} = -\underline{\Lambda}$ (eq. (2.14)) and we cannot satisfy simultaneously both (2.18) and (2.19). **Therefore a two-component $S = \frac{1}{2}$ field cannot be found**.

2.5 Four-component spinor fields

Consider now a four-component function transforming as

$$\psi'(x') = [1 + \underline{\omega} \cdot \underline{\Omega} + \mathrm{i}\underline{\lambda} \cdot \underline{\Lambda}]\psi(x) \tag{2.20}$$

with

$$\underline{\Omega} = \begin{pmatrix} \frac{\mathrm{i}}{2}\underline{\sigma} & \\ & \frac{\mathrm{i}}{2}\underline{\sigma} \end{pmatrix}, \tag{2.21}$$

$$\underline{\Lambda} = \begin{pmatrix} \frac{\mathrm{i}}{2}\underline{\sigma} & \\ & -\frac{\mathrm{i}}{2}\underline{\sigma} \end{pmatrix}. \tag{2.22}$$

These $\underline{\Omega}$ and $\underline{\Lambda}$ satisfy (1.67).

The corresponding angular-momentum operator

$$\underline{S} = -\mathrm{i}\hbar\underline{\Omega} = \frac{\hbar}{2}\begin{pmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{pmatrix} \equiv \frac{\hbar}{2}\underline{\sigma}' \tag{2.23}$$

$$\underline{S}^2 = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \Rightarrow S = \frac{1}{2} \tag{2.24}$$

Introduce the four new matrices

$$\underline{\gamma} = \begin{pmatrix} 0 & \mathrm{i}\underline{\sigma} \\ -\mathrm{i}\underline{\sigma} & 0 \end{pmatrix} \tag{2.25}$$

$$\gamma_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.26}$$

These matrices are *Hermitian*,

$$\gamma_\mu^\dagger = \gamma_\mu \tag{2.27}$$

and their anti-commutators satisfy

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \tag{2.28}$$

Example 2:

$$\{\gamma_4, \gamma_4\} = 2\begin{pmatrix} & & 1 & 0 \\ & & 0 & 1 \\ 1 & 0 & & \\ 0 & 1 & & \end{pmatrix}^2 = 2\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \quad \text{OK}$$

$$-\{\gamma_1,\gamma_2\} = \begin{pmatrix} & & 0 & 1 \\ & & 1 & 0 \\ 0 & -1 & & \\ -1 & 0 & & \end{pmatrix}\begin{pmatrix} & & 0 & -\mathrm{i} \\ & & \mathrm{i} & 0 \\ 0 & \mathrm{i} & & \\ -\mathrm{i} & 0 & & \end{pmatrix} + \begin{pmatrix} & & 0 & -\mathrm{i} \\ & & \mathrm{i} & 0 \\ 0 & \mathrm{i} & & \\ -\mathrm{i} & 0 & & \end{pmatrix}\begin{pmatrix} & & 0 & 1 \\ & & 1 & 0 \\ 0 & -1 & & \\ -1 & 0 & & \end{pmatrix}$$
$$= \begin{pmatrix} -\mathrm{i} & 0 & & \\ 0 & \mathrm{i} & & \\ & & -\mathrm{i} & 0 \\ & & 0 & \mathrm{i} \end{pmatrix} + \begin{pmatrix} \mathrm{i} & 0 & & \\ 0 & -\mathrm{i} & & \\ & & \mathrm{i} & 0 \\ & & 0 & -\mathrm{i} \end{pmatrix} = 0, \quad \text{OK}$$

The present $\underline{\Lambda}$ and $\underline{\Omega}$ in (2.21–2.22) can be expressed as

$$\underline{\Omega} = \frac{1}{2}(\gamma_2\gamma_3, \gamma_3\gamma_1, \gamma_1\gamma_2) \tag{2.29}$$

$$\underline{\Lambda} = \frac{1}{2}(\gamma_1\gamma_4, \gamma_2\gamma_4, \gamma_3\gamma_4). \tag{2.30}$$

In this representation the space inversion is

$$P = \gamma_4. \tag{2.31}$$

This $\underline{\Omega}$, (2.29), commutes with P,

$$[P, \underline{\Omega}] = 0. \tag{2.32}$$

Example 3:

$$[P, \Omega_1] = \frac{1}{2}(\underbrace{\gamma_4\gamma_2\gamma_3}_{\underbrace{-\gamma_2\gamma_4\gamma_3}_{+\gamma_2\gamma_3\gamma_4}} - \gamma_2\gamma_3\gamma_4) = 0.$$

This $\underline{\Lambda}$ anticommutes with P,

$$\{P, \underline{\Lambda}\} = 0 \tag{2.33}$$

Thus the states in this representation can have a well defined parity. The counterparts of our later Dirac matrices are

$$\underline{\alpha} = \mathrm{i}\gamma_4\underline{\gamma} = \begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} \tag{2.34}$$

$$\beta = \gamma_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.35}$$

The γ_μ matrices obey an Algebra.

3 Relativistic wave equations

3.1 Klein-Gordon

The Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{3.1}$$

$$H = \frac{p^2}{2m} + V \tag{3.2}$$

contains first derivatives with respect to time, t, and second derivatives with respect to (xyz). We now wish to handle all four coordinates $(x_1x_2x_3x_4)$ on equal footing.

The **Klein-Gordon** equation was introduced by **Fock** (1926ab), Gordon (1926), Klein (1926-1927), **Kudar** (1926) and Schrödinger (1926). We wish to satisfy

$$E^2 = c^2p^2 + m^2c^4 \tag{3.3}$$

(Compare with (1.40), $m = m_0$).

Taking the square root, we would get

$$H = \sqrt{c^2p^2 + m^2c^4} + V \tag{3.4}$$

but the problem would be to actually take the square root.

Reorganizing, we get

$$(E - V)\psi = \sqrt{c^2p^2 + m^2c^4}\psi \tag{3.5}$$

Squaring (operating twice)

$$\underline{\underline{(E-V)^2\psi}} = (E^2 - 2EV + V^2)\psi = \underline{\underline{(c^2p^2 + m^2c^4)\psi}} \;\big|\; : 2mc^2 \tag{3.6}$$

$$\big[\frac{p^2}{2m} + \frac{1}{2}mc^2 - \frac{1}{2mc^2}(E^2 - 2EV + V^2)\big]\psi = 0 \;\big|\; E = mc^2 + \epsilon$$

$$\big[\frac{p^2}{2m} + \underline{\frac{1}{2}}mc^2 - \frac{1}{2mc^2}(\underline{m^2c^4} + \underline{\underline{2\epsilon mc^2}} + \epsilon^2 - \underline{\underline{2mc^2V}} - 2\epsilon V + V^2)\big]\psi = 0$$

$$\Big[\frac{p^2}{2m} + \underline{\underline{\underline{V - \epsilon}}} - \underbrace{\frac{1}{2mc^2}(\epsilon - V)^2}_{h_m \text{ (mass-velocity) of Pauli}}\Big]\Psi = 0 \tag{3.7}$$

h_m is often expressed as

$$h_m = -\frac{p^4}{8m^3c^2}$$

The external electromagnetic field $\underline{A}, V$ can be included through the *minimal substitution* $\underline{p} \to \underline{p} - \frac{q}{c}\underline{A}$, q being the charge of the particle. For $e \equiv |e|$, (3.6) becomes, for electrons,

$$\underline{\underline{(E-V)^2\psi = \big[c^2(\underline{p} + \frac{e}{c}\underline{A})^2 + m^2c^4\big]\psi}} \tag{3.8}$$

[This is valid in the Gauss-cgs system of units or in atomic units. In SI, use $\underline{p} \to \underline{p} - q\underline{A}$]

For a free particle, replacing

$$E \to \mathrm{i}\hbar\frac{\partial}{\partial t} \tag{3.9}$$

and introducing the **d'Alembert** operator

$$\Box = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}, \tag{3.10}$$

$$\big[-\hbar^2\frac{\partial^2}{\partial t^2} - m^2c^4 + \hbar^2c^2\nabla^2\big]\psi = 0 \quad |:\hbar^2c^2$$

$$\big[\Box - \frac{m^2c^2}{\hbar^2}\big]\psi = 0 \tag{3.11}$$

The equation (3.7) is good for a *pionic* atom, se e.g. **K-C Wang** *et al.* PRA*22*(1980), 1072. It gives a wrong answer for the fine structure of *electronic* atoms[3].

Multiply now (3.11) from left by ψ^*:

$$\begin{aligned} \psi^*\big[\Box - \frac{m^2c^2}{\hbar^2}\big]\psi \quad &= 0, \quad \text{and} \quad | \quad + \\ \big[\Box - \frac{m^2c^2}{\hbar^2}\big]\psi^*\psi \quad &= 0 \qquad\quad | \quad - \end{aligned}$$

$$\psi^*(\Box\psi) - (\Box\psi^*)\psi = 0 \tag{3.12}$$

Using the definition of $\Box$

$$\psi^*(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2})\psi - \big[(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi^*\big]\psi = 0 \tag{3.13}$$

As

$$\psi^*(\psi'') - (\psi^*)''\psi = \big[\psi^*\psi' - (\psi^*)'\psi\big]' \tag{3.14}$$

and

$$\psi^*(\nabla^2\psi) - \psi(\nabla^2\psi^*) = \nabla\cdot\big[\psi^*\nabla\psi - \psi(\nabla\psi^*)\big] \tag{3.15}$$

we get the continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\underline{j} = 0 \tag{3.16}$$

where

$$\rho(\underline{r},t) = \frac{\mathrm{i}\hbar}{2mc^2}\big(\psi^*\frac{\partial\psi}{\partial t} - \frac{\partial(\psi^*)}{\partial t}\psi\big) \tag{3.17}$$

$$\underline{j}(\underline{r},t) = -\frac{\mathrm{i}\hbar}{2m}\big(\psi^*\nabla\psi - (\nabla\psi^*)\psi\big) \tag{3.18}$$

[3]The proof is left as an exercise for the reader

reduce to the right non-relativistic limit (see e.g. Kvantkemi I, p. 20-21). Now the ρ is still real, but not necessarily positive definite. Replace in (3.17) $\mathrm{i}\hbar\frac{\partial}{\partial t} \to E \Rightarrow$

$$\rho = \frac{E}{mc^2}|\psi|^2. \text{ Now, if } E < 0, \ \rho < 0 \tag{3.19}$$

3.2 Dirac

We are searching for an equation of the form

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi = \mathrm{H}\psi \tag{3.20}$$

Because of the first derivative $\frac{\partial}{\partial t}$, we would like to have first derivatives $\frac{\partial}{\partial x}$ etc. as well. Note that the series expansion of the square root in

$$(E - V)\psi = mc^2\sqrt{1 + p^2/c^2}$$

would contain *all* powers of (p^2/c^2).

Suppose now that the Hamiltonian is linear in all $\frac{\partial}{\partial x_\mu}$ and that the wave function Ψ has N components,

$$\Psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix} \tag{3.21}$$

Then the most general free-particle wave equation is

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi_n = \sum_{l=1}^{N}[c\underline{\alpha}_{nl} \cdot \underline{p} + \beta_{nl}mc^2]\Psi_l \tag{3.22}$$

where n runs from $1 \ldots N$, and

$$\underline{\alpha}_{nl} \cdot \underline{p} = -\mathrm{i}\hbar[(\alpha_{nl})_x\frac{\partial}{\partial x} + (\alpha_{nl})_y\frac{\partial}{\partial y} + (\alpha_{nl})_z\frac{\partial}{\partial z}] \tag{3.23}$$

In terms of the $N \times N$ matrices $\underline{\alpha}$ and $\underline{\beta}$,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi = [-\mathrm{i}\hbar c\underline{\alpha} \cdot \nabla + \beta mc^2]\Psi \equiv h_D\Psi \tag{3.24}$$

with the Dirac Hamiltonian

$$h_D = c\underline{\alpha} \cdot \underline{p} + \beta mc^2 \tag{3.25}$$

The D in h_D stands for Dirac.

The components of α are the $N \times N$ matrices ($N \geq 4$, see below) α_x, α_y and α_z. In order for h_D to be Hermitian, $\underline{\alpha}$ and β must be Hermitian:

$$\underline{\alpha}^\dagger = \underline{\alpha}, \quad \beta^\dagger = \beta \tag{3.26}$$

For all points in space-time to be equivalent, $\underline{\alpha}$ and β must be constant and dimensionless. Consequently they commute with $\underline{r}$ and $\underline{p}$.

We still want to satisfy

$$E^2 = c^2p^2 + m^2c^4$$

for all components $\psi_1 \dots \psi_N$:

$$-\hbar^2 \frac{\partial^2}{\partial t^2}\Psi = [-\hbar^2 c^2 \nabla^2 + m^2c^4]\Psi \tag{3.27}$$

Then the Dirac equation

$$h_D\Psi = \mathrm{i}\hbar\frac{\partial}{\partial t}\Psi$$

will connect the different components while every individual component, ψ_i, will satisfy the Klein-Gordon (K-G) equation (3.27).

3.2.1 Why must $N \geq 4$?

Starting from

$$\begin{aligned} \beta\alpha_i &= -\alpha_i\beta \quad |\beta \times \\ \alpha_i &= \underbrace{-\beta}_{A}\underbrace{\alpha_i\beta}_{B} \quad (\text{as } \beta^2 = \mathbf{I}) \end{aligned}$$

we see that

$$\begin{aligned} \text{Tr } AB &= \text{Tr } BA \\ \text{Tr } \alpha_i &= -\text{Tr } (\alpha_i\beta)\beta \quad |\beta^2 = \mathbf{I} \\ &= -\text{Tr } \alpha_i \\ &= 0 \end{aligned}$$

The α_i were ± 1. We have the same number of both $\Rightarrow N$ must be even. On page 17 it was shown that 2 was too small $\Rightarrow N \geq 4$. $\square$

3.2.2 Properties of Dirac matrices

Multiply now

$$(\mathrm{i}\hbar\frac{\partial}{\partial t} - h_D)\Psi = 0$$

from the left by the operator $\mathrm{i}\hbar\frac{\partial}{\partial t} + h_D \Rightarrow$

$$(-\hbar^2\frac{\partial^2}{\partial t^2} - h_D^2)\Psi = 0$$

in other words

$$-\hbar^2\frac{\partial^2}{\partial t^2}\Psi = \big[-\hbar^2c^2\textstyle\sum_{i,j=1}^{3}\frac{1}{2}(\alpha_i\alpha_j+\alpha_j\alpha_i)\frac{\partial^2}{\partial x_i\partial x_j} \\ -\mathrm{i}\hbar c\textstyle\sum_{i=1}^{3}(\beta\alpha_i+\alpha_i\beta)\frac{\partial}{\partial x_i}+m^2c^4\beta^2\big]\Psi \tag{3.28}$$

where we have used the fact that $[\frac{\partial}{\partial x_i},\frac{\partial}{\partial x_j}=0$.

From (3.27),

$$-\hbar^2\frac{\partial^2}{\partial t^2}\Psi = [-\hbar^2c^2\nabla^2+m^2c^4]\Psi$$

$$\begin{cases}\alpha_i\alpha_j+\alpha_j\alpha_i = 2\delta_{ij}\mathbf{I}\\ \beta\alpha_i+\alpha_i\beta = 0\\ \beta^2 = \mathbf{I}\end{cases} \tag{3.29}$$

$\mathbf{I}$ being the $N\times N$ unit matrix. Because $\underline{\alpha}$ and β were Hermitian, their eigenvalues are real. According to (3.29), the squares of these eigenvalues equal 1. Hence the eigenvalues are ±1. Dirac noticed that (3.29) is satisfied by

$$\beta = \begin{pmatrix}\mathbf{I} & \\ & -\mathbf{I}\end{pmatrix} \tag{3.30}$$

$$\underline{\alpha} = \begin{pmatrix}0 & \underline{\sigma}\\ \underline{\sigma} & 0\end{pmatrix} \tag{3.31}$$

where

$$\mathbf{I}=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_2=\begin{pmatrix}0&-\mathrm{i}\\ \mathrm{i}&0\end{pmatrix},\quad \sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix} \tag{3.32}$$

3.3 Some properties of the Dirac equation

3.3.1 Solutions with $\pm E$

For a free particle,

$$[h_D,\underline{p}]=0 \tag{3.33}$$

and hence the two operators have simultaneous eigenfunctions $\Psi(\underline{p},x)$:

$$\begin{aligned}\hat{\underline{p}}\Psi &= \underline{p}\Psi\\ h_D\Psi &= E(\underline{p})\Psi\end{aligned} \tag{3.34}$$

Operating again with h_D,

$$h_D^2\Psi = E(\underline{p})^2\Psi \tag{3.35}$$

with the solutions

$$h_D\Psi_\pm = \pm E(\underline{p})\Psi_\pm, \tag{3.36}$$
$$E(\underline{p}) = \sqrt{c^2p^2 + m^2c^4}$$

corresponding to "electron-like" and "positron-like" solutions, respectively. If only one of them existed, h_D would be a multiple of $\mathbf{I}$, and commute with all matrices.

As (3.29) $\Rightarrow$

$$\begin{aligned}
\{\underline{\alpha}, \beta\} &= 0 \Rightarrow \\
\left[\underline{\alpha}, \beta\right] &= \underline{\alpha}\beta - \beta\underline{\alpha} = -2\beta\underline{\alpha} \Rightarrow \\
\left[h_D, \beta\right] &= -2c\beta\underline{\alpha}\cdot\underline{p} \neq 0 \Rightarrow \\
\exists\beta \text{ so that } [h_d, \beta] \neq 0 &\Rightarrow h_D \neq a\mathbf{I} \Rightarrow \exists\, +E \text{ and } -E
\end{aligned}$$

3.3.2 Inclusion of electromagnetic fields

Use again the minimal substitution $\Rightarrow$

$$h_D = c\underline{\alpha}\cdot\left(\underline{p} - \frac{e}{c}\underline{A}\right) + V + \beta mc^2 \tag{3.37}$$

[In Gauss-cgs. In SI, use $\underline{p} - e\underline{A}$. Now $e < 0$.] The non-relativistic "diamagnetic" term is recovered here in 2^{nd}-order perturbation theory with intermediate positron-like states. Using the closure,

$$\sum_n |\psi_n^-\rangle\langle\psi_n^-| = 1 \tag{3.38}$$
$$\Delta E = 2mc^2$$

we get

$$\begin{aligned}
E^{(2)} &= \sum_n |\langle\psi^+|h_D|\psi_n^-\rangle|^2/(E^+ - E_n^-) \\
&\cong e^2\langle\psi^+|\underline{A}^2|\psi^+\rangle/2mc^2
\end{aligned} \tag{3.39}$$

the usual diamagnetic term. (Recall that $\frac{1}{2m}(\underline{p} - \frac{e}{c}\underline{A})^2 \rightarrow$ the term $\frac{1}{2m}\frac{e^2}{c^2}\underline{A}^2$.)

3.3.3 Free-particle solutions

In atomic units ($e = m_e = \hbar = 4\pi\epsilon_0 = 1$, $c = 137.036$),

$$h_D = c\underline{\alpha}\cdot\underline{p} + c^2\beta \tag{3.40}$$

Using the Ansatz

$$\Psi = u(p)\mathrm{e}^{\mathrm{i}\underline{p}\cdot\underline{r}} \tag{3.41}$$

the Dirac equation $h_D\Psi = E\Psi$ gives

$$\begin{bmatrix} -E+c^2 & 0 & cp_z & c(p_x-\mathrm{i}p_y) \\ 0 & -E+c^2 & c(p_x+\mathrm{i}p_y) & -cp_z \\ cp_z & c(p_x-\mathrm{i}p_y) & -E-c^2 & 0 \\ c(p_x+\mathrm{i}p_y) & -cp_z & 0 & -E-c^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = 0 \tag{3.42}$$

This linear, homogenous equation has solutions if det [] $= 0 \Rightarrow$

$$(E^2 - c^4 - c^2p^2)^2 = 0 \tag{3.43}$$

$$E = \pm c\sqrt{(p^2+c^2)} = \pm E_p \tag{3.44}$$

Denoting $E_p + c^2$ by A, the spinors become (σ_z= the spin):

	(E, σ_z)			
u_i	$(E_p, \frac{1}{2})$	$(E_p, -\frac{1}{2})$	$(-E_p, \frac{1}{2})$	$(-E_p, -\frac{1}{2})$
u_1	1	0	$-cp_z/A$	$-c(p_x - \mathrm{i}p_y)/A$
u_2	0	1	$-c(p_x + \mathrm{i}p_y)/A$	cp_z/A
u_3	cp_z/A	$c(p_x - \mathrm{i}p_y)/A$	1	0
u_4	$c(p_x + \mathrm{i}p_y)/A$	$-cp_z/A$	0	1

(3.45)

The normalised spinors become

$$u(\underline{p}, E_p) = \sqrt{\tfrac{E_p+c^2}{2E_p}} \begin{pmatrix} 1 \cdot \binom{\alpha}{\beta} \\ \frac{c(\underline{\sigma}\cdot\underline{p})}{E_p+c^2}\binom{\alpha}{\beta} \end{pmatrix} = \sqrt{\frac{E_p+c^2}{2E_p}} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{3.46}$$

$$u(\underline{p}, -E_p) = \sqrt{\tfrac{E_p+c^2}{2E_p}} \begin{pmatrix} \frac{-c(\underline{\sigma}\cdot\underline{p})}{E_p+c^2}\binom{\alpha}{\beta} \\ 1 \cdot \binom{\alpha}{\beta} \end{pmatrix} = \sqrt{\frac{E_p+c^2}{2E_p}} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{3.47}$$

where $\binom{\alpha}{\beta}$ is either $\binom{1}{0}$ or $\binom{0}{1}$.

The ratio of the norms for $E_p > 0$,

$$\frac{\psi_2^2}{\psi_1^2} = \frac{c^2p^2}{(E+c^2)^2} \cong \frac{c^2p^2}{(2c^2)^2} = \frac{1}{4}(\frac{p}{mc})^2 \approx \frac{1}{4}(\frac{v}{c})^2 \tag{3.48}$$

for small v/c. Therefore ψ_1 and ψ_2 are called the **large** and **small components**, respectively.

For electrons in light atoms, $v \approx 1$ a.u. and this plane-wave estimate gives

$$\frac{1}{4}(\frac{v}{c})^2 \approx \frac{1}{4}(\frac{1}{137})^2 \approx 10^{-5}$$

3.3.4 Probability density

For Ψ, satisfying

$$\mathrm{i}\hbar\frac{\partial}{\partial t} = -\mathrm{i}\hbar c\underline{\alpha}\cdot\nabla\Psi + \beta mc^2\Psi \quad |\Psi^\dagger\times \tag{3.49}$$

where the Hermitian conjugate

$$\Psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \tag{3.50}$$

satisfies

$$-\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi^\dagger = \mathrm{i}\hbar c(\nabla\Psi^\dagger)\cdot\underline{\alpha} + mc^2\Psi^\dagger\beta \quad |\times\Psi \tag{3.51}$$

Performing the two multiplications and subtracting, we get

$$\mathrm{i}\hbar\frac{\partial}{\partial t}(\Psi^\dagger\Psi) = -\mathrm{i}\hbar c(\Psi^\dagger\underline{\alpha}\cdot\nabla\Psi + \nabla\Psi^\dagger\cdot\underline{\alpha}\Psi) \tag{3.52}$$

$$\frac{1}{c}\frac{\partial}{\partial t}(\Psi^\dagger\Psi) + \nabla\cdot(\Psi^\dagger\underline{\alpha}\Psi) = 0 \tag{3.53}$$

Comparing with the classical continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\underline{j} = 0 \tag{3.54}$$

we identify

$$\rho = \Psi^\dagger\Psi = (\psi_1^\dagger\psi_1 + \psi_2^\dagger\psi_2 + \ldots) \tag{3.55}$$

$$\underline{j} = c\Psi^\dagger\underline{\alpha}\Psi \tag{3.56}$$

From $\underline{j} = \rho\underline{v}$, the velocity operator (local velocity of the Dirac electron)

$$\underline{v} = c\underline{\alpha} \tag{3.57}$$

3.4 The Pauli limit

By substituting $E = mc^2 + \varepsilon$, the Dirac equation becomes

$$\big[c\underline{\alpha}\cdot\underline{p} + \underbrace{(\beta-1)}_{=\begin{pmatrix}0 & 0\\ 0 & -2\mathbf{I}\end{pmatrix}} mc^2 + V\big]\Psi = \varepsilon\Psi \tag{3.58}$$

In terms of the two-spinors

$$\begin{cases} c\underline{\sigma}\cdot\underline{p}\psi_2 + V\psi_1 = \varepsilon\psi_1 \\ c\underline{\sigma}\cdot\underline{p}\psi_1 + (-2mc^2 + V)\psi_2 = \varepsilon\psi_2 \Rightarrow \end{cases} \tag{3.59}$$

$$\psi_2 = \frac{c\underline{\sigma}\cdot\underline{p}}{2mc^2+\varepsilon-V}\psi_1 \tag{3.60}$$

$$\approx \frac{\underline{\sigma}\cdot\underline{p}}{2mc}\psi_1 \tag{3.61}$$

(Note that (3.60) is still exact.) Using the above approximation with $\psi_1 = \psi_{\rm nr}, \langle\psi_1|\psi_1\rangle = 1 \Rightarrow$

$$\begin{aligned}\langle\Psi|\Psi\rangle &= \langle\psi_1|\psi_1\rangle + \langle\psi_2|\psi_2\rangle \\ &= 1 + \langle\psi_2|\psi_2\rangle \qquad |\underline{p}\cdot\underline{\sigma}^\dagger\underline{\sigma}\cdot\underline{p} = \underline{p}^2\mathbf{I} \\ &= 1 + \frac{1}{4m^2c^2}\langle\psi_1|\underline{p}^2|\psi_1\rangle \end{aligned}\tag{3.62}$$

Then the expectation value of h_D becomes

$$\begin{aligned}\varepsilon &= \langle\Psi|h_D|\Psi\rangle/\langle\Psi|\Psi\rangle \\ &= \langle\Psi|c\underline{\alpha}\cdot\underline{p} + (\beta-1)mc^2 + V|\Psi\rangle/\langle\Psi|\Psi\rangle \\ &= \Big[\langle\psi_1|c\underline{\sigma}\cdot\underline{p}|\psi_2\rangle + \langle\psi_2|c\underline{\sigma}\cdot\underline{p}|\psi_1\rangle - 2mc^2\langle\psi_2|\psi_2\rangle \\ &\quad + \langle\psi_1|V|\psi_1\rangle + \langle\psi_2|V|\psi_2\rangle\Big]/\Big[1+\langle\psi_2|\psi_2\rangle\Big] \\ &= \Big[\frac{1}{2m}\langle\psi_1|p^2|\psi_1\rangle + \underline{\frac{1}{2m}\langle\psi_1|p^2|\psi_1\rangle} - \underline{\frac{2mc^2}{4m^2c^2}}\langle\psi_1|p^2|\psi_1\rangle \\ &\quad + \langle\psi_1|V|\psi_1\rangle + \frac{1}{4m^2c^2}\langle\psi_1|\underline{\sigma}\cdot\underline{p}V\underline{\sigma}\cdot\underline{p}|\psi_1\rangle\Big]/\langle\Psi|\Psi\rangle \\ \varepsilon &\cong \Big[\langle\psi_1|\frac{p^2}{2m}|\psi_1\rangle + \langle\psi_1|V|\psi_1\rangle + \frac{1}{4m^2c^2}\langle\psi_1|\underline{\sigma}\cdot\underline{p}V\underline{\sigma}\cdot\underline{p}|\psi_1\rangle\Big] \\ &\quad \times\Big[1 - \frac{1}{4m^2c^2}\langle\psi_1|p^2|\psi_1\rangle\Big] + \mathcal{O}(c^{-4}) \quad \leftarrow 1/c^4 \text{ and higher order terms neglected} \\ &= \varepsilon_{\rm nr} - \frac{1}{8m^3c^2}\langle\psi_1|p^2|\psi_1\rangle\langle\psi_1|p^2|\psi_1\rangle \\ &\quad - \frac{1}{4m^2c^2}\langle\psi_1|V|\psi_1\rangle\langle\psi_1|p^2|\psi_1\rangle \\ &\quad + \frac{1}{4m^2c^2}\langle\psi_1\underline{\sigma}\cdot\underline{p}V\underline{\sigma}\cdot\underline{p}|\psi_1\rangle + \mathcal{O}(c^{-4}) \end{aligned}\tag{3.63}$$

Trick 1:

$$\begin{aligned}(\frac{p^2}{2m}+V)\psi_1 &= \varepsilon_{\rm nr}\psi_1, \\ V\psi_1 &= (\varepsilon_{\rm nr} - \frac{p^2}{2m})\psi_1 \end{aligned}\tag{3.64}$$

$$\begin{aligned}\varepsilon = \varepsilon_{\rm nr} &- \frac{1}{4m^2c^2}\varepsilon_{\rm nr}\langle\psi_1|p^2|\psi_1\rangle \\ &+ \frac{1}{4m^2c^2}\langle\psi_1|\underline{\sigma}\cdot\underline{p}V\underline{\sigma}\cdot\underline{p}|\psi_1\rangle \end{aligned}\tag{3.65}$$

Trick 2:

$$\begin{aligned}
\sigma_i\sigma_j &= \mathrm{i}\delta_{ijk}\sigma_k + \delta_{ij}, && (3.66)\\
\underline{\sigma}\cdot\underline{p}V\underline{\sigma}\cdot\underline{p} &= \sigma_i p_i V \sigma_j p_j \\
&= \mathrm{i}\delta_{ijk}\sigma_k p_i V p_j + p_k V p_k \\
&= \mathrm{i}\delta_{ijk}\sigma_k\big[(p_iV)p_j + V\underbrace{p_ip_j}_{=0}\big] + p_kVp_k \\
&= \hbar\underline{\sigma}\cdot(\nabla V\times\underline{p}) + p_kVp_k && (3.67)
\end{aligned}$$

[Note: $p_ip_j = 0$ as i and j run over all combinations (summation); $p_ip_j - p_jp_i = 0$.]

Here

$$\begin{aligned}
p_kVp_k &= \frac{1}{2}\big[p_kVp_k + p_kVp_k\big] \\
&= \frac{1}{2}\big[Vp^2 + p^2V\big] + \frac{1}{2}\big[(p_kV)p_k - p_k(p_kV)\big] && (3.68)
\end{aligned}$$

where the first term becomes

$$\frac{1}{2}\langle\psi_1|Vp^2 + p^2V|\psi_1\rangle \overset{\mathbf{T1}}{=} \varepsilon_{\mathrm{nr}}\langle\psi_1|p^2|\psi_1\rangle - \frac{1}{2m}\langle\psi_1|p^4|\psi_1\rangle \qquad (3.69)$$

and the second term becomes ($\partial_k \equiv \frac{\partial}{\partial k}$)

$$\begin{aligned}
&\frac{1}{2}\langle\psi_1|(p_kV)p_k - p_k(p_kV)|\psi_1\rangle \\
&= \frac{1}{2}\langle\psi_1|(p_kV)|p_k\psi_1\rangle - \frac{1}{2}\langle p_k\psi_1|(p_kV)|\psi_1\rangle \quad \Big|p_k = -\mathrm{i}\hbar\partial_k \\
&= -\frac{\hbar^2}{2}\Big[\int\psi_1^*(\partial_kV)(\partial_k\psi_1) + \int(\partial_k\psi_1)^*(\partial_kV)\psi_1\Big] \\
&= -\frac{\hbar^2}{2}\Big\{\big[\psi_1^*(\partial_kV)\psi_1\big]_{-\infty}^{\infty} \underline{- \int(\partial_k\psi_1^*)(\partial_kV)\psi_1} - \int\psi_1^*(\partial_k^2V)\psi_1 \\
&\qquad\qquad \underline{+ \int(\partial_k\psi_1)^*(\partial_kV)\psi_1}\Big\} \\
&= \frac{1}{2}\hbar^2\langle\psi_1|\nabla^2V|\psi_1\rangle \quad \Big|\psi(\infty)\to 0 && (3.70)
\end{aligned}$$

Assembling the pieces in (3.65), (3.67), (3.69) and (3.70):

$$\begin{aligned}
\varepsilon &= \varepsilon_{\mathrm{nr}} - \frac{1}{4m^2c^2}\,\underline{\varepsilon_{\mathrm{nr}}\langle\psi_1|p^2|\psi_1\rangle} \\
&\quad + \frac{1}{4m^2c^2}\Big[\langle\psi_1|\hbar\underline{\sigma}\cdot(\nabla V\times\underline{p})|\psi_1\rangle \\
&\qquad + \underline{\varepsilon_{\mathrm{nr}}\langle\psi_1|p^2|\psi_1\rangle} - \frac{1}{2m}\langle\psi_1|p^4|\psi_1\rangle \\
&\qquad + \frac{1}{2}\hbar^2\langle\psi_1|\nabla^2V|\psi_1\rangle\Big] \\
&= \varepsilon_{\mathrm{nr}} + \langle h_m + h_{\mathrm{SO}} + h_{\mathrm{d}}\rangle + \mathcal{O}(c^{-4}) && (3.71)
\end{aligned}$$

$$\begin{cases} \langle h_m\rangle = -\dfrac{1}{8m^3c^2}\langle\psi_1|p^4|\psi_1\rangle & \text{(mass-velocity)} \\ \langle h_{\rm SO}\rangle = \quad \dfrac{\hbar}{4m^2c^2}\langle\psi_1|\underline{\sigma}\cdot(\nabla V\times\underline{p})|\psi_1\rangle & \text{(spin-orbit)} \\ \langle h_{\rm d}\rangle = \quad \dfrac{\hbar^2}{8m^2c^2}\langle\psi_1|\nabla^2 V|\psi_1\rangle & \text{(Darwin)} \end{cases} \tag{3.72}$$

Returning to (3.65), in terms of ψ_2,

$$\varepsilon \cong \varepsilon_{\rm nr} - \varepsilon_{\rm nr} + \langle\psi_2|V|\psi_2\rangle \tag{3.73}$$

For the Coulomb potential of a point charge Z,

$$\nabla^2 V = -Z\nabla^2(\frac{1}{r}) = 4\pi Z\underbrace{\delta(\underline{r})}_{\text{3-D }\delta\text{-func}} \quad \Rightarrow \tag{3.74}$$

$$\begin{aligned} \langle h_{\rm d}\rangle &= \frac{\hbar^2}{8m^2c^2}\langle\psi_1|(4\pi Z)\delta(\underline{r})|\psi_1\rangle \\ &= \frac{\pi\hbar^2 Z}{2m^2c^2}|\psi_1(0)|^2 \end{aligned} \tag{3.75}$$

Example 1: For the 1s state of a hydrogenic atom, in a.u.,

$$\psi_{\rm 1s} = \sqrt{\frac{Z^3}{\pi}}{\rm e}^{-Zr},$$

$$\begin{aligned} \langle h_{\rm d}\rangle &= \frac{\pi Z}{2c^2}\frac{Z^3}{\pi} = \frac{Z^4}{2c^2} \\ \langle h_m\rangle &= -\frac{1}{8m^3c^2}\langle p^4\rangle = -\frac{1}{2mc^2}\langle(\frac{p^2}{2m})^2\rangle \\ &= -\frac{1}{2mc^2}\langle(E-V)^2\rangle \\ &= -\frac{1}{2mc^2}\Big[(\frac{Z^2}{2})^2 - 2(-\frac{Z^2}{2})\langle V\rangle + \langle V^2\rangle\Big] \end{aligned}$$

As

$$\begin{aligned} \langle V\rangle &= \frac{Z^3}{\pi}(4\pi)\int_0^\infty {\rm e}^{-2Zr}\Big(\frac{-Z}{r}\Big)r^2{\rm d}r \\ &= -4Z^4\int_0^\infty {\rm e}^{-2Zr}r{\rm d}r = (-4Z^4)\frac{1!}{(2Z)^2} = -Z^2, \\ \langle V^2\rangle &= +4Z^3\int_0^\infty {\rm e}^{-2Zr}\Big(\frac{Z}{r}\Big)^2 r^2{\rm d}r = 4Z^5\frac{0!}{2Z} = 2Z^4, \\ \langle h_m\rangle &= -\frac{1}{2mc^2}Z^4\Big[\frac{1}{4}-1+2\Big] = -\frac{5}{8c^2}Z^4 \\ \langle h_{\rm SO}\rangle &= 0 \end{aligned}$$

Summing up,

$$E_{\rm rel} = \langle h_m + h_{\rm d} \rangle = \frac{Z^4}{c^2}(-\frac{5}{8} + \frac{4}{8}) = \underline{-\frac{Z^4}{8c^2}} \tag{3.76}$$

3.5 Central fields

We now return to the full Dirac equation with a local potential $V(\underline{r}) = V(r)$. As

$$\begin{aligned} \big[\underline{l}, h_{\rm D}\big] &= \quad {\rm i}c\underline{\alpha} \times \underline{p} \neq 0 \qquad (3.77) \\ \big[\underline{s}, h_{\rm D}\big] &= -{\rm i}c\underline{\alpha} \times \underline{p} \neq 0 \qquad (3.78) \end{aligned}$$

neither $\underline{l}$ or $\underline{s}$ can have simultaneous eigenstates with $h_{\rm D}$ but their sum

$$\underline{j} = \underline{l} + \underline{s} \tag{3.79}$$

can because

$$[\underline{j}, \underline{h_{\rm D}}] = 0 \tag{3.80}$$

Details:

$$\begin{aligned} \underline{l} &= \underline{r} \times \underline{p} \\ \big[\underline{l}, \beta\big] &= 0, \quad \big[\underline{l}, V\big] = 0, \quad \big[\underline{l}, \underline{\alpha} \cdot \underline{p}\big] \neq 0 \end{aligned}$$

Consider the component l_1:

$$[l_1, \underline{\alpha} \cdot \underline{p}] = \alpha_1 \underbrace{[l_1, p_1]}_{=0} + \alpha_2 [l_1, p_2] + \alpha_3 [l_1, p_3]$$

$$\underline{l} = \underline{r} \times \underline{p} = -{\rm i} \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ x & y & z \\ \partial_x & \partial_y & \partial_z \end{vmatrix} \Rightarrow l_1 = -{\rm i}(y\partial_z - z\partial_y)$$

$$\begin{aligned} [l_1, p_2] &= (-{\rm i})^2 [y\partial_z - z\partial_y, \partial_y] \\ &= -\Big(\underline{y\frac{\partial^2}{\partial y \partial z}} \; \underline{\underline{-z\frac{\partial^2}{\partial y^2}}} - \partial_z \; \underline{-y\frac{\partial^2}{\partial y \partial z}} \; \underline{\underline{+z\frac{\partial^2}{\partial y^2}}} \Big) \\ &= \frac{\partial}{\partial z} = {\rm i}p_3, \text{ cycl.} \Rightarrow \\ [l_1, \underline{\alpha} \cdot \underline{p}] = \alpha_2 {\rm i}p_3 - \alpha_3 {\rm i}p_2 &= {\rm i}(\underline{\alpha} \times \underline{p})_1, \text{ cycl.} \Rightarrow \\ [l, \underline{\alpha} \cdot \underline{p}] &= {\rm i}\ \underline{\alpha} \times \underline{p} \quad \square \end{aligned}$$

Similarly for $[\underline{\sigma}, \underline{\alpha}\cdot\underline{p}]$ (=with spin),

$$\begin{aligned}
[\sigma_1, \underline{\alpha}\cdot\underline{p}] &= \begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_1\end{pmatrix}\begin{pmatrix}0 & \underline{\sigma}\cdot\underline{p}\\ \underline{\sigma}\cdot\underline{p} & 0\end{pmatrix} - \begin{pmatrix}0 & \underline{\sigma}\cdot\underline{p}\\ \underline{\sigma}\cdot\underline{p} & 0\end{pmatrix}\begin{pmatrix}\delta_1 & 0\\ 0 & \delta_1\end{pmatrix}\\
&= \begin{pmatrix}0 & \sigma_1\underline{\sigma}\cdot\underline{p}\\ \sigma_1\underline{\sigma}\cdot\underline{p} & 0\end{pmatrix} - \begin{pmatrix}0 & \underline{\sigma}\cdot\underline{p}\sigma_1\\ \underline{\sigma}\cdot\underline{p}\sigma_1 & 0\end{pmatrix}\\
&= \begin{pmatrix}0 & [\sigma_1, \underline{\sigma}\cdot\underline{p}]\\ [\sigma_1, \underline{\sigma}\cdot\underline{p}] & 0\end{pmatrix}
\end{aligned}$$

As $[\sigma_1, \sigma_2] = 2\mathrm{i}\sigma_3$, cycl.,

$$\begin{aligned}
[\sigma_1, \underline{\sigma}\cdot\underline{p}] &= \sigma_1(\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3)\\
&\quad -(\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3)\sigma_1\\
&= p_1 + \sigma_1\sigma_2 p_2 + \sigma_1\sigma_3 p_3\\
&\quad -p_1 - \sigma_2\sigma_1 p_2 - \sigma_3\sigma_1 p_3\\
&= \underbrace{[\sigma_1, \sigma_2]}_{2\mathrm{i}\sigma_3} p_2 - \underbrace{[\sigma_3, \sigma_1]}_{2\mathrm{i}\sigma_2} p_3\\
&= 2\mathrm{i}[\sigma_3 p_2 - \sigma_2 p_3] = 2\mathrm{i}(p_2\sigma_3 - p_3\sigma_2)\\
&= 2\mathrm{i}(\underline{p}\times\underline{\sigma}), \text{ cycl. } \Rightarrow\\
[\underline{\sigma}, \underline{\alpha}\cdot\underline{p}] &= 2\mathrm{i}\underline{p}\times\underline{\alpha} = -2\mathrm{i}\underline{\alpha}\times\underline{p} \quad \square
\end{aligned}$$

Define

$$\underline{s} = \frac{1}{2}\hbar\underline{\sigma} \tag{3.81}$$

$\Rightarrow$ (3.79)

As a further detail, consider the operator

$$K = (V\sigma\cdot\underline{l} + 1)\beta \tag{3.82}$$

We shall show that

$$[K, \underline{j}] = 0 \tag{3.83}$$
$$[K, h_{\mathrm{D}}] = 0 \tag{3.84}$$

In other words, K, h_{D}, $\underline{j}^2$ and j_z can have simultaneous eigenstates.

Theorem *a*:

For two vectors $\underline{A}$ and $\underline{B}$ which commute with $\underline{\sigma}$ but not necessarily with each other,

$$(\underline{\sigma}\cdot\underline{A})(\underline{\sigma}\cdot\underline{B}) = \underline{A}\cdot\underline{B} + \mathrm{i}\underline{\sigma}\cdot(\underline{A}\times\underline{B}) \tag{3.85}$$

Proof:
The $\underline{\sigma}$ matrices simultaneously satisfy

$$\left.\begin{aligned}[\sigma_i,\sigma_k] = \sigma_i\sigma_k - \sigma_k\sigma_i = 2\mathrm{i}\delta_{ikl}\delta_l \\ \{\sigma_i,\sigma_k\} = \sigma_i\sigma_k + \sigma_k\sigma_i = 2\delta_{ik}\end{aligned}\right\} \Rightarrow$$

$$\sigma_i^2 = 1 \qquad (i=k) \tag{3.86}$$

$$\sigma_i\sigma_k = \mathrm{i}\delta_{ikl}\sigma_l \quad (i\neq k) \tag{3.87}$$

$$\begin{aligned}(\underline{\sigma}\cdot\underline{A})(\underline{\sigma}\cdot\underline{B}) &= \sigma_i A_i \sigma_j B_j \\ &\overset{[\underline{\sigma},\underline{A}]=0}{=} \underbrace{\sigma_i\sigma_i}_{=1} A_iB_i + \sum_{i\neq j}\sigma_i\sigma_j A_iB_j \\ &= \underline{A}\cdot\underline{B} + \sum_{i<j}\underbrace{\sigma_i\sigma_j}_{\mathrm{i}\delta_{ijk}\delta_k}\underbrace{(A_iB_j - A_jB_i)}_{(\underline{A}\times\underline{B})_k} \\ &= \underline{A}\cdot\underline{B} + \mathrm{i}\underline{\sigma}\cdot\underline{A}\times\underline{B} \quad \square\end{aligned} \tag{3.88}$$

The equation holds for both two- and four-component $\underline{\sigma}$

Theorem *b*:

$$(\underline{\alpha}\cdot\underline{A})(\underline{\alpha}\cdot\underline{B}) = (\underline{\sigma}\cdot\underline{A})(\underline{\sigma}\cdot\underline{B}) \tag{3.89}$$

Proof:

$$\begin{aligned}\begin{pmatrix}0 & \underline{\sigma}\cdot\underline{A}\\ \underline{\sigma}\cdot\underline{A} & 0\end{pmatrix}\begin{pmatrix}0 & \underline{\sigma}\cdot\underline{B}\\ \underline{\sigma}\cdot\underline{B} & 0\end{pmatrix} &= \begin{pmatrix}\underline{\sigma}\cdot\underline{A}\,\underline{\sigma}\cdot\underline{B} & 0\\ 0 & \underline{\sigma}\cdot\underline{A}\,\underline{\sigma}\cdot\underline{B}\end{pmatrix} \\ &= (\underline{\sigma}^4\cdot\underline{A})(\underline{\sigma}^4\cdot\underline{B}) \quad \square\end{aligned}$$

$\underline{\sigma}^4$ denotes a four dimensional σ.

Theorem *c*:

$$(\underline{\alpha}\cdot\underline{A})(\underline{\alpha}\cdot\underline{B}) = -\gamma_5\underline{A}\cdot\underline{B} + \mathrm{i}\underline{\alpha}\cdot(\underline{A}\times\underline{B}), \tag{3.90}$$

where

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4. \tag{3.91}$$

Proof:
We use the representation

$$\underline{\gamma} = \mathrm{i}\underline{\alpha}\beta = \begin{pmatrix} 0 & -\mathrm{i}\underline{\sigma} \\ \mathrm{i}\underline{\sigma} & 0 \end{pmatrix}$$
$$\gamma_5 = \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Then, recalling that $\mathrm{i}^3 = -\mathrm{i}$,

$$\begin{aligned} \gamma_5 &= -\mathrm{i}\underbrace{\begin{pmatrix} 0 & -\sigma_1 \\ \sigma_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix}}\ \underbrace{\begin{pmatrix} 0 & -\sigma_3 \\ \sigma_3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \\ &= -\mathrm{i}\begin{pmatrix} -\sigma_1\sigma_2 & 0 \\ 0 & -\sigma_1\sigma_2 \end{pmatrix} \qquad \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix} \\ &= \quad \mathrm{i}\begin{pmatrix} 0 & \sigma_1\sigma_2\sigma_3 \\ \sigma_1\sigma_2\sigma_3 & 0 \end{pmatrix} \qquad \Big| \sigma_1\sigma_2 = \mathrm{i}\sigma_3 \\ &= -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned} \tag{3.92}$$

Then

$$\begin{cases} \gamma_5\underline{\alpha} = -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} = -\begin{pmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{pmatrix} = -\underline{\sigma}^4 \\ \underline{\alpha}\gamma_5 = -\begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -\begin{pmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{pmatrix} = -\underline{\sigma}^4, \end{cases} \tag{3.93}$$

$$\begin{cases} \gamma_5\underline{\sigma}^4 = -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{pmatrix} = -\begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} = -\underline{\alpha} \\ \underline{\sigma}^4\gamma_5 = -\begin{pmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -\begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} = -\underline{\alpha}. \end{cases} \tag{3.94}$$

Multiplying **Theorem *a***, (3.85), from left by $-\gamma_5$:

$$(-\underbrace{\gamma_5\underline{\sigma}^4}_{-\underline{\alpha}}\cdot\underline{A})(\underline{\sigma}\cdot\underline{B}) = -\gamma_5\underline{A}\cdot\underline{B} + \mathrm{i}(\underbrace{-\gamma_5\underline{\sigma}^4}_{\underline{\alpha}})\cdot\underline{A}\times\underline{B} \quad \square$$

Theorem *d*:
Because

$$\begin{aligned} \underline{l} &= \underline{r}\times\underline{p} \quad (\perp \underline{p}), \\ \underline{p}\cdot\underline{l} &= \underline{l}\cdot\underline{p} = 0. \end{aligned} \tag{3.95}$$

Furthermore,

$$\underline{l}\times\underline{p} + \underline{p}\times\underline{l} = 2\mathrm{i}\underline{p} \tag{3.96}$$

Proof:

$$\begin{aligned}
\underline{A}\times\underline{B} &= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \\
\underline{A}\times\underline{B}+\underline{B}\times\underline{A} &= \hat{\imath}\big([A_y,B_z]-[A_z,B_y]\big) \\
&\quad +\hat{\jmath}\big(-[A_z,B_x]+[A_x,B_z]\big) \\
&\quad +\hat{k}\big([A_x,B_y]-[A_y,B_x]\big)
\end{aligned}$$

$$\begin{aligned}
(\underline{l}\times\underline{p}+\underline{p}\times\underline{l})_x &= [l_y,p_z]-[l_z,p_y] \\
&= -\big[(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z})(\frac{\partial}{\partial z})-(\frac{\partial}{\partial z})(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z})\big] \\
&\quad +\big[(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x})(\frac{\partial}{\partial y})-(\frac{\partial}{\partial y})(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x})\big] \\
&= \frac{\partial}{\partial x}+\frac{\partial}{\partial x}=2\mathrm{i}p_x, \quad \text{cycl.}\,\square
\end{aligned}$$

Then the anticommutator

$$\begin{aligned}
\{\underline{\sigma}\cdot\underline{l},\underline{\alpha}\cdot\underline{p}\} &= (\underline{\sigma}\cdot\underline{l})(\underline{\alpha}\cdot\underline{p})+(\underline{\alpha}\cdot\underline{p})(\underline{\sigma}\cdot\underline{l}) \\
&\overset{(3.90)}{=} -\gamma_5[\underbrace{\underline{l}\cdot\underline{p}}_{=0}+\underbrace{\underline{p}\cdot\underline{l}}_{=0}+\mathrm{i}\underline{\sigma}^4\cdot(\underbrace{\underline{l}\times\underline{p}+\underline{p}\times\underline{l}}_{2\mathrm{i}\underline{p}})] \\
&= -2\underline{\alpha}\cdot\underline{p} \qquad (3.97)
\end{aligned}$$

Theorem *e*:

$$[K,\underline{\alpha}\cdot\underline{p}]=[\beta\underline{\sigma}\cdot\underline{l},\underline{\alpha}\cdot\underline{p}]+[\beta,\underline{\alpha}\cdot\underline{p}]$$

Because

$$\begin{aligned}
-\underline{\alpha}\beta &= \beta\underline{\alpha}, \\
-\underline{\alpha}\cdot\underline{p}\beta\underline{\sigma}\cdot\underline{l} &= \beta(\underline{\alpha}\cdot\underline{p})\underline{\sigma}\cdot\underline{l}\Rightarrow \\
[K,\underline{\alpha}\cdot\underline{p}] &= \beta\underbrace{\{\underline{\sigma}\cdot l,\underline{\alpha}\cdot\underline{p}\}}_{-2\underline{\alpha}\cdot\underline{p}}+2\beta\underline{\alpha}\cdot\underline{p}=0 \qquad (3.98)
\end{aligned}$$

Theorem *f*:
As

$$[\beta,\beta]=0, \quad [K,\beta]=0. \qquad (3.99)$$

Theorem e+f $\Rightarrow$

$$[K, h_{\mathrm{D}}] = 0 \tag{3.100}$$

It remains to show that $[K, \underline{j}] = 0$.

Theorem *g*:

From theorem a, (3.85), taking $\underline{B}$ as an arbitrary vector, commuting with $\underline{A}$,

$$(\underline{\sigma} \cdot \underline{A})\underline{\sigma} = \underline{A} + \mathrm{i}\underline{\sigma} \times \underline{A}, \tag{3.101}$$

where we used

$$\underline{\sigma} \cdot \underline{A} \times \underline{B} = \underline{\sigma} \times \underline{A} \cdot \underline{B}.$$

Similarly, letting $\underline{A}$ be arbitrary,

$$\underline{\sigma}(\underline{\sigma} \cdot \underline{B}) = \underline{B} + \mathrm{i}\underline{B} \times \underline{\sigma}. \tag{3.102}$$

Letting then $\underline{A} = \underline{B} = \underline{l}$ and subtracting (3.101)-(3.102),

$$\begin{aligned} [\underline{\sigma} \cdot \underline{l}, \underline{\sigma}] &= \mathrm{i}\underline{\sigma} \times \underline{l} - \mathrm{i}\underline{l} \times \underline{\sigma} \\ &= 2\mathrm{i}\underline{\sigma} \times \underline{l} \end{aligned} \tag{3.103}$$

Theorem *h*:

Let $\underline{A}$ commute with $\underline{l}$. Then

$$\begin{aligned} [\underline{A} \cdot \underline{l}, \underline{l}] &= \sum_{i=1}^{3} \sum_{j=1}^{3} [A_i l_i l_j \hat{x}_j - l_j \hat{x}_j A_i l_i] \\ &= \sum_{i,j} A_i \hat{x}_j \underbrace{[l_i l_j - l_j l_i]}_{\delta_{ijk}\mathrm{i}l_k = -\mathrm{i}\delta_{jik}l_k} \\ &= -\mathrm{i}\underline{A} \times \underline{l} \end{aligned} \tag{3.104}$$

Letting $\underline{A} = \underline{\sigma}$,

$$[\underline{\sigma} \cdot \underline{l}, \underline{l}] = -\mathrm{i}\underline{\sigma} \times \underline{l} \tag{3.105}$$

Theorem *i*:

The desired commutator (note: $[K, \beta] = 0$)

$$\begin{aligned} [K, \underline{j}] &= [\beta\underline{\sigma} \cdot \underline{l}, \underline{l} + \frac{1}{2}\underline{\sigma}] \\ &= \beta[\underline{\sigma} \cdot \underline{l}, \underline{l}] + \frac{1}{2}\beta[\underline{\sigma} \cdot \underline{l}, \underline{\sigma}] \\ &= \beta(-\mathrm{i}\underline{\sigma} \times \underline{l}) + \frac{1}{2}\beta(2\mathrm{i}\underline{\sigma} \times \underline{l}) = 0 \quad \square \end{aligned} \tag{3.106}$$

Summarizing,

$$\boxed{[K, \underline{j}] = [K, h_{\rm D}] = [\underline{j}, h_{\rm D}] = 0} \tag{3.107}$$

Denote the eigenvalues of K by $-\kappa$:

$$K\chi_\kappa^m = -\kappa\chi_\kappa^m, \tag{3.108}$$

where the $\underline{\sigma}$ (in K) are 2-component ones, and the 2-component spinors

$$\chi_\kappa^m \equiv |lsjm\rangle \tag{3.109}$$

The quantum number κ carries both j and l:

$$\kappa = \begin{cases} -l-1, & j = l + \dfrac{1}{2} \quad \text{(state '}l\text{')} \\ \quad l, & j = l - \dfrac{1}{2} \quad \text{(state '}l^*\text{' or '}\bar{l}\text{')}. \end{cases} \tag{3.110}$$

The lowest values are

$$\begin{array}{cccccccccc} \kappa & = & \text{-1} & 1 & \text{-2} & 2 & \text{-3} & 3 & \text{-4} & \dots \\ & & s_{1/2} & p_{1/2} & p_{3/2} & d_{3/2} & d_{5/2} & f_{5/2} & f_{7/2} & \dots \\ & & \text{s} & \text{p}^* & \text{p} & \text{d}^* & \text{d} & \text{f}^* & \text{f} & \dots \end{array} \tag{3.111}$$

From

$$\begin{aligned} \underline{j} &= \underline{l} + \underline{s}, \\ \underline{j}^2 &= \underline{l}^2 + 2\underline{l}\cdot\underline{s} + \underline{s}^2 \Rightarrow \\ \underline{l}\cdot\underline{s} &\equiv \frac{1}{2}[\underline{j}^2 - \underline{l}^2 - \underline{s}^2] \end{aligned} \tag{3.112}$$

Consequently

$$\begin{aligned} K &= \underline{\sigma}\cdot\underline{l} + 1 \\ &= \underline{j}^2 - \underline{l}^2 - \underline{s}^2 + 1 \\ K\chi_\kappa^m &= K|lsjm\rangle = [j(j+1) - l(l+1) - \underbrace{s(s+1)}_{\frac{3}{4}} + 1]\chi_\kappa^m \\ &= [(j+\frac{1}{2})^2 - l(l+1)]\chi_\kappa^m \\ &= -\kappa\chi_\kappa^m, \end{aligned}$$

whence

$$\kappa = l(l+1) - (j+\frac{1}{2})^2, \tag{3.113}$$

as can be verified from Table (3.111).

For the 4-component case we introduce the Ansatz

$$\Psi = \Psi_\kappa^m = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix} g(r)\chi_\kappa^m \\ \mathrm{i}f(r)\chi_{-\kappa}^m\end{pmatrix} = \frac{1}{r}\begin{pmatrix} P(r)\chi_\kappa^m \\ \mathrm{i}Q(r)\chi_{-\kappa}^m\end{pmatrix} \tag{3.114}$$

for which, indeed

$$K\Psi_\kappa^m = \begin{pmatrix}\underline{\sigma}\cdot\underline{l}+1 & 0\\ 0 & -\underline{\sigma}\cdot\underline{l}-1\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = -\kappa\Psi_\kappa^m. \tag{3.115}$$

This Ψ is an eigenfunction to $h_\mathrm{D}, K, \underline{j}^2$ and j_z.

3.5.1 The radial Part

From

$$\begin{aligned}\underline{A}\times(\underline{B}\times\underline{C}) &= (\underline{A}\cdot\underline{C})\underline{B} - (\underline{A}\cdot\underline{B})\underline{C},\\ (\underline{A}\cdot\underline{B})\underline{C} &= (\underline{A}\cdot\underline{C})\underline{B} - \underline{A}\times(\underline{B}\times\underline{C}),\end{aligned} \tag{3.116}$$

with

$$\begin{aligned}\underline{A} &= \underline{B} = \underline{\hat{r}} = (\hat{\imath},\hat{\jmath},\hat{k}), \quad \underline{C} = \nabla,\\ \nabla &= \underline{\hat{r}}(\hat{r}\cdot\nabla) - \underline{\hat{r}}\times(\underline{\hat{r}}\times\nabla)\\ &= \hat{e}_r\frac{\partial}{\partial r} - \mathrm{i}\frac{\hat{e}_r}{r}\times\underline{l}.\end{aligned} \tag{3.117}$$

Here $\hat{e}_r$ is the unit vector along the radius (and $\underline{l} = \underline{r}\times\underline{p}$). Consequently the kinetic energy term

$$\underline{\alpha}\cdot\underline{p} = -\mathrm{i}\alpha_r\frac{\partial}{\partial r} - \frac{1}{r}\underline{\alpha}\cdot\hat{r}\times\underline{l}. \tag{3.118}$$

Denoting

$$\underline{\sigma} = \underline{\sigma}^4 = \begin{pmatrix}\underline{\sigma} & 0\\ 0 & \underline{\sigma}\end{pmatrix}$$

and using

$$\underline{\alpha}\cdot\underline{A}\,\underline{\alpha}\cdot\underline{B} = \underline{A}\cdot\underline{B} + \mathrm{i}\underline{\sigma}\cdot\underline{A}\times\underline{B} \qquad \text{[3.88-3.89]}$$

for $\underline{A} = \underline{\hat{r}}$ and $\underline{B} = \underline{l}$,

$$\underbrace{(\underline{\alpha}\cdot\underline{\hat{r}})}_{\alpha_r}(\underline{\alpha}\cdot\underline{l}) = \underbrace{\underline{\hat{r}}\cdot\underline{l}}_{=0} + \mathrm{i}\underline{\sigma}\cdot\underline{\hat{r}}\times\underline{l} \tag{3.119}$$

($\underline{\hat{r}}\cdot\underline{l} = 0$ as the angular momentum has no radial components.) Multiplying by $\mathrm{i}\gamma_5$ and recalling that $\underline{\alpha} = -\gamma_5\underline{\sigma}$,

$$\begin{aligned}\mathrm{i}\gamma_5\alpha_r(\underline{\alpha}\cdot\underline{l}) &= -\gamma_5\underline{\sigma}\cdot\underline{r}\times\underline{l}\\ \mathrm{i}\gamma_5\sigma_r(\underline{\sigma}\cdot\underline{l}) &= \underline{\alpha}\cdot\underline{r}\times\underline{l}.\end{aligned} \tag{3.120}$$

$$\begin{aligned}\underline{\alpha}\cdot\underline{p} &= \mathrm{i}\gamma_5\sigma_r\frac{\partial}{\partial r} - \frac{1}{r}\mathrm{i}\gamma_5\sigma_r(\underline{\sigma}\cdot\underline{l}) \\ &= \mathrm{i}\gamma_5\sigma_r[\frac{\partial}{\partial r} - \frac{1}{r}\underline{\sigma}\cdot\underline{l}]. \end{aligned} \tag{3.121}$$

Thus the Dirac Hamiltonian

$$\begin{aligned}h_{\mathrm{D}} &= c\underline{\alpha}\cdot\underline{p} + \beta mc^2 + V(r) \\ &= \mathrm{i}c\gamma_5\sigma_r\Big[\frac{\partial}{\partial r} - \frac{1}{r}\underline{\sigma}\cdot\underline{l}\Big] + \beta mc^2 + V(r). \end{aligned} \tag{3.122}$$

with

$$\begin{aligned}K = \beta(\underline{\sigma}\cdot\underline{l}+1), \quad \beta^2 = 1 \Rightarrow \\ \underline{\sigma}\cdot\underline{l} = \beta K - 1, \end{aligned} \tag{3.123}$$

$$\boxed{h_{\mathrm{D}} = \mathrm{i}c\gamma_5\sigma_r\Big[\tfrac{\partial}{\partial r} + \tfrac{1}{r} - \tfrac{\beta K}{r}\Big] + \beta mc^2 + V(r)} \tag{3.124}$$

Theorem *a*:

$$\sigma_r^2 = 1 \tag{3.125}$$

(like in the cartesian case, $\sigma_i^2 = 1$)

Proof:

$$\sigma_r = \underline{\sigma}\cdot\hat{r} = \sum_i \sigma_i\frac{1}{r}(\sigma_x x + \sigma_y y + \sigma_z z)$$

$$\sigma_r^2 = frac1r^2\Big[\sum_i \underbrace{\sigma_i^2}_{=1} x_i^2 + \sum_{i>j} x_i x_j \underbrace{(\sigma_i\sigma_j + \sigma_j\sigma_i)}_{=0}\Big] = 1 \quad \square$$

Theorem *b*:

$$\{K, \sigma_r\} = 0 \tag{3.126}$$

Proof:

$$\begin{aligned}K &= \beta(\underline{\sigma}\cdot\underline{l}+1) \\ [\beta, \underline{\sigma}\cdot\underline{l}] &= 0 \\ [\beta, \sigma_r] &= 0\end{aligned}$$

(3.85) $\Rightarrow$

$$\left.\begin{aligned}(\underline{\sigma}\cdot\underline{l})(\underline{\sigma}\cdot\hat{\underline{r}}) &= \underline{l}\cdot\hat{\underline{r}} + \mathrm{i}\underline{\sigma}\cdot(\underline{l}\times\hat{\underline{r}})\\ (\underline{\sigma}\cdot\hat{\underline{r}})(\underline{\sigma}\cdot\underline{l}) &= \hat{\underline{r}}\cdot\underline{l} + \mathrm{i}\underline{\sigma}\cdot(\hat{\underline{r}}\times\underline{l})\end{aligned}\right\}\Rightarrow$$

$$\begin{aligned}\{K,\sigma_r\} &= \{\beta(\underline{\sigma}\cdot\underline{l}+1),\underline{\sigma}\cdot\hat{\underline{r}}\}\\ &= \beta\mathrm{i}\underline{\sigma}\cdot\underbrace{(\underline{l}\times\hat{\underline{r}}+\hat{\underline{r}}\times\underline{l})}_{2\mathrm{i}\underline{\hat{r}}}+2\beta\underline{\sigma}\cdot\hat{\underline{r}}\\ &= 0 \quad \square\end{aligned}$$

Theorem *c*:

$$\sigma_r\chi^m_\kappa = -\chi^m_{-\kappa} \tag{3.127}$$

Proof:

$$\begin{aligned}K(\sigma_r\chi^m_\kappa) &= -\sigma_r(K\chi^m_\kappa)\\ &= \kappa(\sigma_r\chi^m_\kappa)\end{aligned} \tag{3.128}$$

Thus the eigenvalue of $\sigma_r\chi^m_\kappa$ is κ, while

$$K\chi^m_\kappa = -\kappa\chi^m_\kappa \qquad [3.108]$$

whence

$$\sigma_r\chi^m_\kappa = -\chi^m_{-\kappa} \quad \square$$

Recalling now that

$$\gamma_5 = -\begin{pmatrix}0&1\\1&0\end{pmatrix}, \quad \beta = \begin{pmatrix}1&0\\0&-1\end{pmatrix},$$

(3.124) becomes

$$\left[\mathrm{i}c(-)\begin{pmatrix}0&1\\1&0\end{pmatrix}\sigma_r\left(\frac{\partial}{\partial r}+\frac{1}{r}-\frac{1}{r}\begin{pmatrix}1&0\\0&-1\end{pmatrix}K^4\right)+\begin{pmatrix}1&0\\0&-1\end{pmatrix}mc^2+V-E\right]\begin{pmatrix}g\chi^m_\kappa\\ \mathrm{i}f\chi^m_{-\kappa}\end{pmatrix}=0.$$

where K^4 denotes the four-dimensional K-operator. Recalling

$$K^4\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix} = (-\kappa)\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix},$$

$$-\mathrm{i}c\begin{pmatrix}0&1\\1&0\end{pmatrix}\sigma_r\left[\Big(\frac{\partial}{\partial r}+\frac{1}{r}\Big)\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}-\frac{1}{r}\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}\right]$$
$$+mc^2\begin{pmatrix}g\chi_\kappa\\ -\mathrm{i}f\chi_{-\kappa}\end{pmatrix}+(V-E)\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}=0.$$

Then operate with σ_r using (3.127):

$$-\mathrm{i}c\begin{pmatrix}0&1\\1&0\end{pmatrix}\left[-\Big(\frac{\partial}{\partial r}+\frac{1}{r}\Big)\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}+\frac{1}{r}\begin{pmatrix}g\chi_{-\kappa}\\ \mathrm{i}f\chi_{\kappa}\end{pmatrix}\right]$$
$$+mc^2\begin{pmatrix}g\chi_\kappa\\ -\mathrm{i}f\chi_{-\kappa}\end{pmatrix}+(V-E)\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}=0,$$

$$\mathrm{i}c\Big(\frac{\partial}{\partial r}+\frac{1}{r}\Big)\begin{pmatrix}\mathrm{i}f\chi_\kappa\\ g\chi_{-\kappa}\end{pmatrix}-\frac{\mathrm{i}c}{r}\begin{pmatrix}\mathrm{i}f\kappa\chi_\kappa\\ -g\kappa\chi_{-\kappa}\end{pmatrix}$$
$$+mc^2\begin{pmatrix}g\chi_\kappa\\ -\mathrm{i}f\chi_{-\kappa}\end{pmatrix}+(V-E)\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}=0.$$

For the two components

$$\mathrm{i}c\left(\frac{\partial}{\partial r}+(1\{\mp\}\kappa)\frac{1}{r}\right)\begin{pmatrix}\mathrm{i}f\chi_\kappa\\ g\chi_{-\kappa}\end{pmatrix}+\Big[\{\pm\}mc^2+V-E\Big]\begin{pmatrix}g\chi_\kappa\\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}=0 \quad (3.129)$$

Dividing by the angular part and i we get

$$\begin{cases}-c\Big(\dfrac{\partial}{\partial r}+\dfrac{1-\kappa}{r}\Big)f+(mc^2+V-E)g &=0\\ c\Big(\dfrac{\partial}{\partial r}+\dfrac{1+\kappa}{r}\Big)g+(-mc^2+V-E)f &=0\end{cases} \quad (3.130)$$

Introducing

$$\begin{cases}P=rg\\ Q=rf\end{cases}, \quad (3.131)$$

$$\begin{cases}\dfrac{\mathrm{d}P}{\mathrm{d}r}=g+r\dfrac{\mathrm{d}g}{\mathrm{d}r}\\ \dfrac{\mathrm{d}Q}{\mathrm{d}r}=f+r\dfrac{\mathrm{d}f}{\mathrm{d}r}\end{cases}\Rightarrow\begin{cases}\dfrac{\mathrm{d}g}{\mathrm{d}r}=\dfrac{1}{r}\dfrac{\mathrm{d}P}{\mathrm{d}r}-\dfrac{P}{r^2}\\ \dfrac{\mathrm{d}f}{\mathrm{d}r}=\dfrac{1}{r}\dfrac{\mathrm{d}Q}{\mathrm{d}r}-\dfrac{Q}{r^2}\end{cases} \quad (3.132)$$

$$\begin{cases}\dfrac{1}{r}\Big(\dfrac{\mathrm{d}Q}{\mathrm{d}r}-\dfrac{Q}{r}+(1-\kappa)\dfrac{Q}{r}\Big)+\Big(-mc+\dfrac{E-V}{c}\Big)\dfrac{P}{r}=0\\ \dfrac{1}{r}\Big(\dfrac{\mathrm{d}P}{\mathrm{d}r}-\dfrac{P}{r}+(1+\kappa)\dfrac{P}{r}\Big)+\Big(-mc+\dfrac{V-E}{c}\Big)\dfrac{Q}{r}=0\end{cases} \quad (3.133)$$

$$\boxed{\begin{cases} \dfrac{\mathrm{d}P}{\mathrm{d}r} + \kappa\dfrac{P}{r} + \Big(-mc + \dfrac{V-E}{c}\Big)Q = 0 \\ \dfrac{\mathrm{d}Q}{\mathrm{d}r} - \kappa\dfrac{Q}{r} + \Big(-mc + \dfrac{E-V}{c}\Big)P = 0 \end{cases}} \tag{3.134}$$

3.5.2 Non-relativistic limit

Let the potential $|V| \ll E \approx mc^2$. For the $E > 0$ states

$$-mc + \frac{V-E}{c} \approx -2mc \Rightarrow$$

$$Q \approx \frac{1}{2mc}\Big(\frac{\mathrm{d}P}{\mathrm{d}r} + \kappa\frac{P}{r}\Big) \Rightarrow \tag{3.135}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}r} = \frac{1}{2mc}\Big(P'' + \frac{\kappa}{r}P' - \frac{\kappa}{r^2}P\Big) \tag{3.136}$$

Denoting $E = mc^2 + \varepsilon$, for bound states $\varepsilon < 0$,

$$P'' + \underline{\frac{\kappa}{r}P'} - \frac{\kappa}{r^2}P - \frac{\kappa}{r} \times \Big(\underline{P'} + \kappa\frac{P}{r}\Big) + (2mc)\Big(-mc + \frac{mc^2 + \varepsilon - V}{c}\Big)P = 0.$$

$$P'' - \frac{\kappa(\kappa+1)}{r^2}P + 2m(\varepsilon - V)P = 0. \tag{3.137}$$

$$\kappa = \begin{cases} -l-1 \\ l \end{cases} \Rightarrow \kappa(\kappa+1) = \begin{cases} (-l-1)(-l) \\ l(l+1) \end{cases} = l(l+1) \tag{3.138}$$

$\Rightarrow$

$$P'' - \frac{l(l+1)}{r^2}P + 2m(\varepsilon - V)P = 0, \tag{3.139}$$

the usual non-relativistic equation.

3.6 The Dirac-Coulomb Problem

(Darwin 1928, Gordon 1928). Consider

$$V = -\frac{Z}{r} \tag{3.140}$$

For normalized bound states

$$\int_0^\infty r^2 \mathrm{d}r(f^2 + g^2) = \int_0^\infty \mathrm{d}r(P^2 + Q^2) = 1 \tag{3.141}$$

(3.134) $\Rightarrow$

$$\begin{cases} \dfrac{\mathrm{d}P}{\mathrm{d}r} + \kappa\dfrac{P}{r} + \Big(-mc - \dfrac{Z}{c}\dfrac{1}{r} - \dfrac{E}{c}\Big)Q = 0 \\ \dfrac{\mathrm{d}Q}{\mathrm{d}r} - \kappa\dfrac{Q}{r} + \Big(-mc + \dfrac{Z}{c}\dfrac{1}{r} + \dfrac{E}{c}\Big)P = 0 \end{cases} \tag{3.142}$$

So far $\hbar = 1$. Let now $m = c = 1$, as well. Then the unit charge $e = \alpha = 1/137.036$. Furthermore denote

$$\lambda = \sqrt{(}1 - E^2) \tag{3.143}$$

$$x = 2\lambda r \tag{3.144}$$

and introduce the new functions ϕ_1 and ϕ_2 for which

$$\begin{cases} P = \sqrt{1+E}\,\mathrm{e}^{-\lambda r}(\phi_1 + \phi_2) \\ Q = \sqrt{1-E}\,\mathrm{e}^{-\lambda r}(\phi_1 - \phi_2) \end{cases} \tag{3.145}$$

Then

$$\begin{cases} \dfrac{\mathrm{d}P}{\mathrm{d}r} = \sqrt{1+E}\,\mathrm{e}^{-\lambda r}\Big[-\lambda(\phi_1+\phi_2) + \dfrac{\mathrm{d}\phi_1}{\mathrm{d}r} + \dfrac{\mathrm{d}\phi_2}{\mathrm{d}r}\Big] \\ \dfrac{\mathrm{d}Q}{\mathrm{d}r} = \sqrt{1-E}\,\mathrm{e}^{-\lambda r}\Big[-\lambda(\phi_1-\phi_2) + \dfrac{\mathrm{d}\phi_1}{\mathrm{d}r} - \dfrac{\mathrm{d}\phi_2}{\mathrm{d}r}\Big] \end{cases} \tag{3.146}$$

Straightforward algebra on Eq. (3.142) gives:

$$\begin{cases} \sqrt{1+E}\,\mathrm{e}^{-\lambda r}\Big[-\lambda(\phi_1+\phi_2) + \dfrac{\mathrm{d}\phi_1}{\mathrm{d}r} + \dfrac{\mathrm{d}\phi_2}{\mathrm{d}r}\Big] = \\ \qquad -\dfrac{\kappa}{r}\sqrt{1+E}\,\mathrm{e}^{-\lambda r}(\phi_1+\phi_2) + \Big[E + \dfrac{Z}{r} + 1\Big]\sqrt{1-E}\,\mathrm{e}^{-\lambda r}(\phi_1-\phi_2) \end{cases}$$

Dividing the top equation by $\sqrt{1+E}\mathrm{e}^{-\lambda r}$ and the bottom one by $\sqrt{1-E}\mathrm{e}^{-\lambda r}$ and noting that $\frac{\partial}{\partial r} = 2\lambda\frac{\partial}{\partial x}$ we get

$$\begin{cases} -\lambda(\phi_1+\phi_2) + 2\lambda\Big(\dfrac{\mathrm{d}\phi_1}{\mathrm{d}x} + \dfrac{\mathrm{d}\phi_2}{\mathrm{d}x}\Big) = \\ \qquad -\dfrac{\kappa}{r}(\phi_1+\phi_2) + \Big[E + \dfrac{2\lambda Z}{x} + 1\Big]\sqrt{\dfrac{1-E}{1+E}}(\phi_1-\phi_2) \\ -\lambda(\phi_1-\phi_2) + 2\lambda\Big(\dfrac{\mathrm{d}\phi_1}{\mathrm{d}x} - \dfrac{\mathrm{d}\phi_2}{\mathrm{d}x}\Big) = \\ \qquad \dfrac{\kappa}{r}(\phi_1-\phi_2) - \Big[E + \dfrac{2\lambda Z}{x} - 1\Big]\sqrt{\dfrac{1+E}{1-E}}(\phi_1+\phi_2) \end{cases}$$

Dividing now by 2λ, gives

$$\begin{cases} \phi_1' + \phi_2' = \dfrac{1}{2}(\phi_1 + \phi_2) - \dfrac{\kappa}{x}(\phi_1 + \phi_2) \\ \qquad\qquad + \Big[E + \dfrac{2\lambda Z}{x} + 1\Big] \dfrac{1}{2\sqrt{(1+E)(1-E)}} \sqrt{\dfrac{1-E}{1+E}} (\phi_1 - \phi_2) \\ \phi_1' - \phi_2' = \dfrac{1}{2}(\phi_1 - \phi_2) + \dfrac{\kappa}{x}(\phi_1 - \phi_2) \\ \qquad\qquad - \Big[E + \dfrac{2\lambda Z}{x} - 1\Big] \dfrac{1}{2\sqrt{(1+E)(1-E)}} \sqrt{\dfrac{1+E}{1-E}} (\phi_1 + \phi_2) \end{cases}$$

Now, remembering the definition of λ, (3.143), the above can be reorganized to

$$\begin{cases} \phi_1' + \phi_2' = \phi_1 \Big[\dfrac{1}{2} - \dfrac{\kappa}{x} + \dfrac{E + \frac{2\lambda Z}{x} + 1}{2(1+E)}\Big] \\ \qquad\qquad + \phi_2 \Big[\dfrac{1}{2} - \dfrac{\kappa}{x} - \dfrac{E + \frac{2\lambda Z}{x} + 1}{2(1+E)}\Big] \\ \phi_1' - \phi_2' = \phi_1 \Big[\dfrac{1}{2} + \dfrac{\kappa}{x} - \dfrac{E + \frac{2\lambda Z}{x} - 1}{2(1-E)}\Big] \\ \qquad\qquad + \phi_2 \Big[-\dfrac{1}{2} - \dfrac{\kappa}{x} - \dfrac{E + \frac{2\lambda Z}{x} - 1}{2(1-E)}\Big] \end{cases}$$

Writing down the sum and difference of the above { (sum at top, difference at bottom):

$$\begin{cases} 2\phi_1' = \phi_1 \Big[1 + \Big(E + \dfrac{2\lambda Z}{x}\Big)\dfrac{1}{2}\Big(\dfrac{1}{1+E} - \dfrac{1}{1-E}\Big) + \dfrac{1}{2}\Big(\dfrac{1}{1+E} + \dfrac{1}{1-E}\Big)\Big] \\ \qquad + \phi_2 \Big[-\dfrac{2\kappa}{x} + \Big(E + \dfrac{2\lambda Z}{x}\Big)\dfrac{-1}{2}\Big(\dfrac{1}{1+E} + \dfrac{1}{1-E}\Big) - \dfrac{1}{2}\Big(\dfrac{1}{1+E} - \dfrac{1}{1-E}\Big)\Big] \\ 2\phi_2' = \phi_1 \Big[-\dfrac{2\kappa}{x} + \Big(E + \dfrac{2\lambda Z}{x}\Big)\dfrac{1}{2}\Big(\dfrac{1}{1+E} + \dfrac{1}{1-E}\Big) + \dfrac{1}{2}\Big(\dfrac{1}{1+E} - \dfrac{1}{1-E}\Big)\Big] \\ \qquad + \phi_2 \Big[1 - \Big(E + \dfrac{2\lambda Z}{x}\Big)\dfrac{1}{2}\Big(\dfrac{1}{1+E} - \dfrac{1}{1-E}\Big) - \dfrac{1}{2}\Big(\dfrac{1}{1+E} + \dfrac{1}{1-E}\Big)\Big] \end{cases} \tag{3.147}$$

As

$$\begin{aligned} \frac{1}{1+E} - \frac{1}{1-E} &= \frac{-2E}{1-E^2} = \frac{-2E}{\lambda^2} \\ \frac{1}{1+E} + \frac{1}{1-E} &= \frac{2}{1-E^2} = \frac{2}{\lambda^2}, \end{aligned} \tag{3.148}$$

$$\begin{cases} 2\phi_1' = & \phi_1\left[1+\left(E+\dfrac{2\lambda Z}{x}\right)\left(\dfrac{-E}{\lambda^2}\right)+\dfrac{1}{\lambda^2}\right] \\ & +\phi_2\left[-\dfrac{2\kappa}{x}-\left(E+\dfrac{2\lambda Z}{x}\right)\dfrac{1}{\lambda^2}+\dfrac{E}{\lambda^2}\right] \\ 2\phi_2' = & \phi_1\left[-\dfrac{2\kappa}{x}+\left(E+\dfrac{2\lambda Z}{x}\right)\dfrac{1}{\lambda^2}-\dfrac{E}{\lambda^2}\right] \\ & +\phi_2\left[1-\left(E+\dfrac{2\lambda Z}{x}\right)\left(\dfrac{-E}{\lambda^2}\right)-\dfrac{1}{\lambda^2}\right] \end{cases} \tag{3.149}$$

$$\begin{cases} \phi_1' = \phi_1\left[1-\dfrac{ZE}{x\lambda}\right]+\phi_2\left[-\dfrac{\kappa}{x}-\dfrac{Z}{x\lambda}\right] \\ \phi_2' = \phi_1\left[-\dfrac{\kappa}{x}+\dfrac{Z}{x\lambda}\right]+\phi_2\left[\dfrac{ZE}{x\lambda}\right] \end{cases} \tag{3.150}$$

Introducing the series

$$\begin{cases} \phi_1 = x^\gamma \displaystyle\sum_{m=0}^{\infty} \alpha_m x^m \\ \phi_2 = x^\gamma \displaystyle\sum_{m=0}^{\infty} \beta_m x^m \end{cases} \tag{3.151}$$

$$\begin{cases} \phi_1' = x^\gamma \displaystyle\sum_{m=0}^{\infty} (\gamma+m)\alpha_m x^{m-1} \\ \quad = x^\gamma \displaystyle\sum_{m=0}^{\infty} \left[\alpha_m x^m\left(1-\dfrac{ZE}{x\lambda}\right)-\beta_m x^m\left(\dfrac{\kappa}{x}+\dfrac{Z}{x\lambda}\right)\right] \\ \phi_2' = x^\gamma \displaystyle\sum_{m=0}^{\infty} (\gamma+m)\beta_m x^{m-1} \\ \quad = x^\gamma \displaystyle\sum_{m=0}^{\infty} \left[\alpha_m x^m\left(-\dfrac{\kappa}{x}+\dfrac{Z}{x\lambda}\right)+\beta_m x^m \dfrac{ZE}{x\lambda}\right] \end{cases} \tag{3.152}$$

The coefficients for every power, $x^{\gamma+m-1}$, must vanish for the expansion to be true:

$$\begin{cases} \alpha_m(\gamma+m) = \alpha_{m-1} - \alpha_m \dfrac{ZE}{\lambda} - \beta_m\left(\kappa+\dfrac{Z}{\lambda}\right) \\ \beta_m(\gamma+m) = \alpha_m\left(-\kappa+\dfrac{Z}{\lambda}\right)+\beta_m\dfrac{ZE}{\lambda}. \end{cases} \tag{3.153}$$

If α_{m-1} is known, these equations gives (α_m, β_m) For $m=0$ ($\alpha_{-1}=0$ is assumed),

$$\begin{cases} \gamma\alpha_0 = \alpha_0\left(-\dfrac{ZE}{\lambda}\right)-\beta_0\left(\kappa+\dfrac{Z}{\lambda}\right) \\ \gamma\beta_0 = \alpha_0\left(-\kappa+\dfrac{Z}{\lambda}\right)+\beta_0\dfrac{ZE}{\lambda}. \end{cases} \tag{3.154}$$

This linear, homogenous pair has solutions if

$$\begin{vmatrix} \gamma + ZE/\lambda & \kappa + Z/\lambda \\ \kappa - Z/\lambda & \gamma - ZE/\lambda \end{vmatrix} = 0 \tag{3.155}$$

$$\gamma^2 - \left(\frac{ZE}{\lambda}\right)^2 - \kappa^2 + \left(\frac{Z}{\lambda}\right)^2 = 0$$

$$\gamma_2 + \left(\frac{Z}{\lambda}\right)^2 (1 - E^2) - \kappa^2 = 0$$

$$\gamma^2 = \kappa^2 - Z^2. \tag{3.156}$$

It turn out that

$$\gamma = +\sqrt{\kappa^2 - Z^2} \tag{3.157}$$

is the acceptable solution and

$$\gamma = -\sqrt{\kappa^2 - Z^2}$$

the irregular one (which can't be normalized). Recall that Z included the proton charge $e_{\rm p} = \alpha = 1/137.036$. For integer Z',

$$\gamma = +\sqrt{\kappa^2 - (\alpha Z')^2} \tag{3.158}$$

The Eq. (3.153b) $\Rightarrow$

$$\frac{\beta_m}{\alpha_m} = \frac{-\kappa + Z/\lambda}{m + \gamma - ZE/\lambda} = \frac{\kappa - Z/\lambda}{n' - m}, \tag{3.159}$$

In terms of a new quantity

$$n' = \frac{ZE}{\lambda} - \gamma. \tag{3.160}$$

A substitution to (3.153a) gives

$$\alpha_m \left[\gamma + m + \frac{ZE}{\lambda} + \left(\kappa + \frac{Z}{\lambda}\right)\frac{\kappa - Z/\lambda}{n' - m}\right] = \alpha_{m-1} \tag{3.161}$$

Here

$$\begin{aligned}
&\left[\gamma+m+\frac{ZE}{\lambda}+\Big(\kappa+\frac{Z}{\lambda}\Big)\frac{\kappa-Z/\lambda}{n'-m}\right]\\
&=\gamma+m+\frac{ZE}{\lambda}+\frac{\kappa^2-(Z/\lambda)^2}{n'-m}\\
&=\frac{1}{n'-m}\left[\Big(\gamma+m+\frac{ZE}{\lambda}\Big)\Big(-\gamma-m+\frac{ZE}{\lambda}\Big)+\kappa^2-(Z/\lambda)^2\right]\\
&=\frac{1}{n'-m}\left[-\Big(\gamma^2+2\gamma m+m^2\Big)+(ZE/\lambda)^2+\kappa^2-(Z/\lambda)^2\right]\\
&=\frac{1}{n'-m}\left[-\kappa^2+Z^2-2\gamma m-m^2+\kappa^2+Z^2\underbrace{\frac{E^2-1}{\lambda^2}}_{=-1}\right]\\
&=-m\frac{2\gamma+m}{n'-m}\Rightarrow
\end{aligned}$$

$$\alpha_m=-\frac{n'-m}{m(2\gamma+m)}\alpha_{m-1}. \tag{3.162}$$

$$\begin{aligned}
\alpha_m&=(-)^m\frac{(n'-m)\dots(n'-1)}{m!(2\gamma+m)\dots(2\gamma+1)}\alpha_0\\
&=\frac{(1-n')(2-n')\dots(m-n')}{m!(2\gamma+1)\dots(2\gamma+m)}\alpha_0.
\end{aligned} \tag{3.163}$$

(3.159) $\Rightarrow$

$$\left.\begin{aligned}\frac{\beta_m}{\alpha_m}&=\frac{\kappa-Z/\lambda}{n'-m}\\ \frac{\beta_{m-1}}{\alpha_{m-1}}&=\frac{\kappa-Z/\lambda}{n'-m+1}\end{aligned}\right\}\Rightarrow$$

$$\begin{aligned}
\frac{\beta_m}{\beta_{m-1}}&=\frac{n'-m+1}{n'-m}\frac{\alpha_m}{\alpha_{m-1}}\Rightarrow\\
\beta_m&=(-)^m\frac{(n'-m+1)\dots(n')}{m!(2\gamma+1)\dots(2\gamma+m)}\beta_0.
\end{aligned} \tag{3.164}$$

(3.154b) $\Rightarrow$

$$\begin{aligned}
\alpha_0\Big(-\kappa+\frac{Z}{\lambda}\Big)&=\beta_0\Big(\gamma-\frac{ZE}{\lambda}\Big)\Rightarrow\\
\frac{\alpha_0}{\beta_0}=\frac{\gamma-ZE/\lambda}{-\kappa+Z/\lambda}&=\frac{n'}{\kappa-Z/\lambda}.
\end{aligned} \tag{3.165}$$

Comparing the coefficients (3.163)-(3.164) with those of a confluent hypergeometrical function,

$$ {}_1F_1(a;c;x) = 1 + \frac{a}{c}x + \frac{a(a+1)}{c(c+1)}\frac{x^2}{2!} + \cdots , \tag{3.166} $$

$$ \begin{cases} \phi_1 = \alpha_0 x^\gamma {}_1F_1(1-n';2\gamma+1;x) \\ \phi_2 = \beta_0 x^\gamma {}_1F_1(-n';2\gamma+1;x) \\ \quad = \dfrac{\kappa - Z/\lambda}{n'}\alpha_0 x^\gamma {}_1F_1(-n';2\gamma+1;x). \end{cases} \tag{3.167} $$

For $x \to \infty$, ${}_1F_1(a;c;x) \to \mathrm{e}^x$ because

$$ \lim_{m\to\infty}\frac{\alpha_m}{\alpha_{m-1}} = \lim_{m\to\infty}\frac{-n'+m}{m(2\gamma+m)} \to \frac{1}{m}. \tag{3.168} $$

Therefore the series must terminate, That happens for ϕ_2 with

$$ n' = 0, 1, 2, \ldots \tag{3.169} $$

($n' = 0$ is the ground state, $n' = 1$ the 1^{st} excited state, etc.) Thus n' is the number of nodes in ϕ_2 (or in g)

For $n'1 = 0$, the coefficients of ϕ_1 have

$$ \begin{aligned} \alpha_m &= \frac{(1-n')\ldots(m-n')}{m!(2\gamma+1)\ldots(2\gamma+m)}\alpha_0 \\ &= \frac{1}{m!(2\gamma+1)\ldots(2\gamma+m)}\alpha_0, \end{aligned} $$

which converges.

Introduce the principal quantum number

$$ n = n' + |\kappa|. \tag{3.170} $$

Then the definition of n', (3.160), gives

$$ \begin{aligned} \frac{ZE_{n\kappa}}{\lambda} &= n' + \gamma \\ E^2 = \frac{\lambda^2}{Z^2}(n'+\gamma)^2 &= \frac{(1-E^2)(n'+\gamma)^2}{Z^2} \Rightarrow \\ E^2\Big[1 + \frac{(n'+\gamma)^2}{Z^2}\Big] &= \frac{(n'+\gamma)^2}{Z^2} \Rightarrow \end{aligned} $$

$$ \boxed{E_{n\kappa} = \Big[1 + \tfrac{Z^2}{(n'+\gamma)^2}\Big]^{-1/2}} \tag{3.171} $$

The κ occurs only in $\gamma = \sqrt{\kappa^2 - Z^2}$, $+|\kappa|$ and $-|\kappa|$ are degenerate for the same n! For instance for $n = 2$,

$$E_{2s_{1/2}} = E_{2p_{1/2}} < E_{2p_{3/2}}. \tag{3.172}$$

(Note though that with QED accounted for $E_{2s_{1/2}} > E_{2p_{1/2}}$!) An expansion in powers of Z^2 gives (we still have $c = 1$)

$$E_{n\kappa} - 1 = \varepsilon_{n\kappa} = -\frac{1}{2}\frac{Z^2}{n^2} + \frac{Z^4}{n^3}\Big[\frac{3}{8n} - \frac{1}{2|\kappa|}\Big] + \mathcal{O}(Z^6) \tag{3.173}$$

Qualitative conclusions:

a. Largest stabilization for $|\kappa| = 1$.

b. States with same l (and n) but different j are *spin-orbit* split.

c. For $|\kappa| = 1$, both f and g diverge at the origin like $r^{\gamma-1}$. They remain normalized. The isomer or isotope shifts, proportional to $\rho(0)$, grow by a factor of 13 for U.

d. The radial density, $P^2 + Q^2$, has no nodes.

e. The radial electron density suffers a relativistic contraction.

f. Normalization is no problem

g. Continuum, also positron-like can be solved.

h. Integrals, properties, solvable.

i. Literature: See Table 2.3 of RTAM I-III (1986, 1993, 2000)

3.7 Virial theorems

Literature: See RTAM I-III, Table 2.5.

Fock (1930b), Gupta (1932), Rose & Welton (1952), March (1953), Kim (1967), McKinley (1971), M. Brack (1983).

3.7.1 Non-relativistic case

Let

$$\underline{r} \to \underline{r}' = \lambda \underline{r} \quad |\underline{r} = \frac{\underline{r}'}{\lambda} \tag{3.174}$$

$$H = T + V = -\frac{1}{2}\nabla^2 + Cr^n$$

$$\to \lambda^2\Big(-\frac{1}{2}\nabla'^2\Big) + C\lambda^{-n}(r')^n \tag{3.175}$$

$$E(\lambda) = \lambda^2\langle T\rangle + \lambda^{-n}\langle V\rangle \tag{3.176}$$

$$\frac{\partial E}{\partial \lambda} = \Big[2\lambda\langle T\rangle - n\lambda^{-n-1}\langle V\rangle\Big]_{\lambda=1} = 0 \tag{3.177}$$

$\Rightarrow$

$$\boxed{2\langle T\rangle = n\langle V\rangle} \quad \text{Coulomb: } n=1 \tag{3.178}$$

Because $E(\lambda)$ has to be a minimum at $\lambda = 1$,

$$\Big(\frac{\partial^2 E}{\partial\lambda^2}\Big)_{\lambda=1} = \Big[2\langle T\rangle + n(n+1)\lambda^{-n-2}\langle V\rangle\Big]_{\lambda=1} > 0 \tag{3.179}$$

$\Rightarrow$

$$\begin{gathered} n\langle V\rangle + n(n+1)\langle V\rangle > 0, \quad \langle V\rangle < 0 \\ (2+n)n\langle V\rangle > 0 \end{gathered}$$

$$\boxed{-2 < n < 0} \tag{3.180}$$

3.7.2 Dirac

$$\begin{aligned} h_{\mathrm{D}}\Phi &= E\Phi && (3.181)\\ h_{\mathrm{D}} &= c\underline{\alpha}\cdot\underline{p}0\beta mc^2 + V(r), && (3.182)\\ V(r) &= Cr^n && (3.183) \end{aligned}$$

Bound states considered:

$$\Phi = \begin{pmatrix}\phi_A\\ \phi_B\end{pmatrix}, \quad \langle\phi|\phi\rangle = 1, \quad \langle A|A\rangle + \langle B|B\rangle = 1 \tag{3.184}$$

Again, let $\underline{r} \to \underline{r}' = \lambda\underline{r}$, $\mathrm{d}\underline{r} \to \lambda^{-1}\mathrm{d}r'$.

$$\begin{aligned} \Phi(\underline{r}) &\to \lambda^{3/2}\Phi(\underline{r}') && (3.185)\\ h_{\mathrm{D}} &\to c\lambda\underline{\alpha}\cdot\underline{p} + \beta mc^2 + \lambda^{-n}V(r). && (3.186) \end{aligned}$$

$$\begin{cases} \langle\Phi'|c\underline{\alpha}\cdot\underline{p}|\Phi'\rangle = \lambda\langle\Phi|c\underline{\alpha}\cdot\underline{p}|\Phi\rangle \\ \langle\Phi'|\beta mc^2|\Phi'\rangle = mc^2(N_A - N_B) \\ \langle\Phi'|r^{-n}|\Phi'\rangle = \lambda^{-n}\langle\Phi|r^{-n}|\Phi\rangle \end{cases} \tag{3.187}$$

$$\begin{aligned} E(\lambda) &= \lambda\langle c\underline{\alpha}\cdot\underline{p}\rangle + \lambda^{-n}\langle V\rangle + mc^2(N_A - N_B) && (3.188)\\ \Big(\frac{\partial E}{\partial\lambda}\Big)_{\lambda=1} &= 0 \quad \Rightarrow \boxed{\langle T\rangle = n\langle V\rangle} && (3.189) \end{aligned}$$

For $n = -1$,

$$\langle T\rangle + \langle V\rangle = c\langle\underline{\alpha}\cdot\underline{p}\rangle + \langle V\rangle = 0 \tag{3.190}$$

$$\Big(\frac{\partial^2 E}{\partial \lambda^2}\Big)_{\lambda=1} \Rightarrow$$

$$\big[n(n+1)\lambda^{-n-2}\langle V\rangle\big]_{\lambda=1} > 0 \tag{3.191}$$

$$\langle V\rangle < 0 \Rightarrow n(n+1) < 0 \Rightarrow$$
$$-1 < \ n \ < 0. \tag{3.192}$$

For $n = -1$,

$$E = \langle T\rangle + \langle V\rangle + mc^2(N_A - N_B)$$
$$E = mc^2(N_A - N_B). \tag{3.193}$$

$$\boxed{E = mc^2\langle\beta\rangle} \tag{3.194}$$

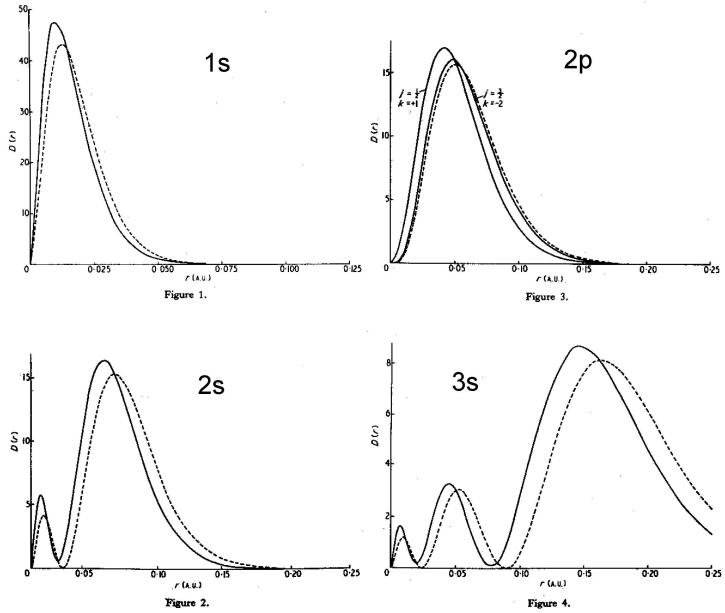

Figure 3.1: Radial electron densities for the 1s, 2s, 2p and 3s states in a hydrogen-like atom with $Z = 80$. The dotted curve denotes the non-relativistic density and the solid curve the relativistic density. Reproduced with permission from V.M. Burke and I.P. Grant, *Proc. Phys. Soc. (London)* **90** (1967) 297-314.

From Figure (3.1) it is seen that the relativistic contraction is of the order for the states 1s, 2s, $2p_{1/2}$ and 3p, whereas the contraction of $2p_{3/2}$ is considerably lesser. The figures 2 and 4 (inside (3.1)) show that 2s and 3s are procentually contracted as much as 1s is. The same conclusion can be drawn for all s-electrons up to the valence shell in a many-electron atom. The p-shells also contract in the same way, but not as much.

4 Dirac-Fock

4.1 The energy-expression

Subtracting the rest energy mc^2 from E, the one-electron Dirac Hamiltonian

$$h_i = c\underline{\alpha}_i \cdot \underline{p}_i + mc^2(\beta_i - I) - \frac{Z}{r_i} \tag{4.1}$$

and the **Dirac-Fock Hamiltonian**

$$H = \sum_i h_i + \sum_{i<j} r_{ij}^{-1} \tag{4.2}$$

The total wave function is the **Slater determinant**

$$\Psi = (N!)^{-1/2}|\psi_1(1)\psi_2(2)\dots\psi_N(N)|, \tag{4.3}$$

where the one-electron wave functions

$$\psi_{n\kappa m} = \begin{pmatrix} r^{-1}P_{n\kappa}(r)\chi_\kappa^m(\vartheta,\phi,\sigma) \\ \mathrm{i}r^{-1}Q_{n\kappa}(r)\chi_{-\kappa}^m(\vartheta,\phi,\sigma) \end{pmatrix} \tag{4.4}$$

Then the total energy

$$\begin{aligned} E &= \langle\Psi|H|\Psi\rangle \\ &= \sum_i \langle\psi_i|h|\psi_i\rangle + \sum_{i<j}\Big[\langle\psi_i\psi_j|\frac{1}{r_{12}}|\psi_i\psi_j\rangle - \langle\psi_i\psi_j|\frac{1}{r_{12}}|\psi_j\psi_i\rangle\Big] \\ &= \sum_i I_i + \sum_{i<j}\big[J_{ij} - K_{ij}\big], \end{aligned} \tag{4.5}$$

in terms of the one-electron integrals I, Coulomb integrals J and exchange integrals K.

Eq. (3.129) $\Rightarrow$

$$\begin{aligned} I &= \langle\psi|h|\psi\rangle \\ &= \langle(g\chi_\kappa, -\mathrm{i}f\chi_{-\kappa}\Big| \times \\ &\Big[\mathrm{i}c\big(\frac{\partial}{\partial r} + \frac{1+\{\mp n\}}{r}\big)\begin{pmatrix} \mathrm{i}f\chi_\kappa \\ g\chi_{-\kappa}\end{pmatrix} + \big(\begin{Bmatrix} 0 \\ -2mc^2\end{Bmatrix} + V\big)\begin{pmatrix} g\chi_\kappa \\ \mathrm{i}f\chi_{-\kappa}\end{pmatrix}\Big]\rangle \end{aligned}$$

$$
\begin{aligned}
=&\int_0^\infty r^2\mathrm{d}r\Big\{g\Big[c\Big(-\frac{\partial}{\partial r}+\frac{\kappa-1}{r}\Big)f+Vg\Big]\\
&\qquad +f\Big[c\Big(\ \frac{\partial}{\partial r}+\frac{\kappa+1}{r}\Big)g+(-2mc^2+V)f\Big]\Big\}\\
=&\int_0^\infty \mathrm{d}r\Big\{P\Big[c\Big(-\frac{\mathrm{d}Q}{\mathrm{d}r}+\frac{f}{r}-\frac{f}{r}+\frac{\kappa}{r}Q\Big)+VP\Big]\\
&\qquad +Q\Big[c\Big(\frac{\mathrm{d}P}{\mathrm{d}r}-\frac{q}{r}+\frac{g}{r}+\frac{\kappa}{r}P\Big)+(-2mc^2+V)Q\Big]\Big\}\\
=&\int_0^\infty \mathrm{d}r\Big\{c\,Q\Big(\frac{\mathrm{d}P}{\mathrm{d}r}+\frac{\kappa}{r}P\Big)-cP\Big(\frac{\mathrm{d}Q}{\mathrm{d}r}-\frac{\kappa}{r}Q\Big)+V(r)(P^2+Q^2)-2mc^2Q^2\Big\}=0
\end{aligned}
\tag{4.6}
$$

$$
g=\frac{P}{r},\quad g'=\frac{1}{r}P'-\frac{P}{r^2}=\frac{1}{r}\frac{\mathrm{d}P}{\mathrm{d}r}-\frac{1}{r}g \tag{4.7}
$$

We see that $I_{n\kappa}$ is independent of m.

Consider next the general two-electron integral

$$
\begin{aligned}
\langle\psi_a\psi_b|\frac{1}{r_{12}}|\psi_c\psi_d\rangle &= \langle ab|v|cd\rangle\\
&= \int \mathrm{d}\underline{r}_1\int \mathrm{d}\underline{r}_2\psi_a^*(1)\psi_b^*(2)\frac{1}{r_{12}}\psi_c(1)\psi_d(2)
\end{aligned}
\tag{4.8}
$$

using the expansion

$$
\begin{aligned}
\frac{1}{r_{12}} &= \sum_{k=0}^\infty \frac{r_<^k}{r_>^{k+1}}C^k(1)\cdot C^k(2)\\
&= \sum_k \frac{r_<^k}{r_>^{k+1}}\sum_{q=-k}^{k}(-1)^q C_q^k(1)C_{-q}^k(2).
\end{aligned}
\tag{4.9}
$$

Here

$$
C_q^k=\sqrt{\frac{4\pi}{2k+1}}Y_k^q,\quad -k\le q\le k. \tag{4.10}
$$

Then the angular parts

$$
\begin{aligned}
\langle\kappa m|C_q^k|\kappa' m'\rangle &= (-)^{m+\frac{1}{2}}\Big[(2j+1)(2j'+1)\Big]^{1/2}\times\\
&\qquad \begin{pmatrix} j & k & j'\\ \frac{1}{2} & 0 & -\frac{1}{2}\end{pmatrix}\begin{pmatrix} j & k & j'\\ -m & q & m'\end{pmatrix}\delta(m,q+m')\\
&\equiv d^k(j'm',jm)
\end{aligned}
\tag{4.11}
$$

and the radial **Slater integrals**

$$\begin{aligned} R^k(abcd) &= \int_0^\infty \mathrm{d}r_1 \int_0^\infty \mathrm{d}r_2 \frac{r_<^k}{r_>^{k+1}} \big[P_a(1)P_c(1) + Q_a(1)Q_c(1)\big] \\ &\quad \times \big[P_b(2)P_d(2) + Q_b(2)Q_d(2)\big] \\ &= \int_0^\infty \frac{1}{r_2} Y_k(ac, r_2)\big[P_b(2)P_d(2) + Q_b(2)Q_d(2)\big]\mathrm{d}r_2 \end{aligned} \tag{4.12}$$

where

$$Y_k(ac, r_2) = r_2 \int_0^\infty \frac{r_<^k}{r_>^{k+1}} \big[P_a(1)P_c(1) + Q_a(1)Q_c(1)\big]\mathrm{d}r_1 \tag{4.13}$$

$\Rightarrow$

$$\langle ab|v|cd\rangle = \sum_{kq} (-)^q R^k(abcd)\langle a|C_q^k|c\rangle\langle b|C_{-q}^k|d\rangle. \tag{4.14}$$

The **3-j symbols** fulfil

$$|j_3 m_3\rangle = (2j_3+1)^{1/2} \sum_{m_1,m_2} (-)^{-j_1+j_2-m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} |j_1 m_1\rangle |j_2 m_2\rangle. \tag{4.15}$$

In other words,

$$C(j_1 j_2 j_3; m_1 m_2 m_3) = (-)^{-j_1+j_2-m_3}(2j_3+1)^{1/2} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}, \tag{4.16}$$

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-)^{j_1-j_2-m_3}(2j_3+1)^{-1/2} C(j_1 j_2 j_3; m_1 m_2 - m_3). \tag{4.17}$$

The 3-j symbols are tabulated by **Rotenberg** *et al.* For their *symmetries*, see p. 107 of Kvantkemi I. They can be calculated by a simple program.

In the special case $j_1 = l$, $j_2 = \frac{1}{2}$, $C(l\frac{1}{2}j; m_l m_s m)$ is:

	m_s	
j	$\frac{1}{2}$	$-\frac{1}{2}$
$l+\frac{1}{2}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$	$\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}$
$l-\frac{1}{2}$	$-\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$

$m_l = m - m_s$.

Returning to Eq. (4.11), if the quantum numbers m and m' are fixed, only one value $q = m - m'$ survives.

From

$$\langle ab|v|cd\rangle = \sum_{kq}(-)^q R(abcd)\langle a|C^k_q|c\rangle\langle b|C^k_{-q}|d\rangle, \qquad [4.14]$$

$$\begin{aligned}\langle ab|v|ab\rangle &= \sum_{kq}(-)^q R(abab)\langle a|C^k_q|a\rangle\langle b|C^k_{-q}|b\rangle \\ &\overset{q=0}{=} \sum_k d^k(j_a m_a, j_a m_a) d^k(j_b m_b, j_b m_b) F^k(a,b) \\ &= \sum_k a^k(j_a m_a, j_b m_b) F^k(a,b), \end{aligned} \qquad (4.18)$$

where

$$a^k(j_a m_a, j_b m_b) = d^k(j_a m_a, j_a m_a) d^k(j_b m_b, j_b m_b) \qquad (4.19)$$

and

$$F^k(a,b) = R^k(abab) \qquad (4.20)$$

This was the **Coulomb integral**. The **exchange part**

$$\begin{aligned}\langle ab|v|ba\rangle &= \sum_{kq}(-)^q R(abba)\langle a|C^k_q|b\rangle\langle b|C^k_{-q}|a\rangle \\ &= \sum_k b^k(j_a m_a, j_b m_b) G^k(a,b) \end{aligned} \qquad (4.21)$$

with

$$b^k(jm; j'm') = [d^k(jm; j'm')]^2 \qquad (4.22)$$

and

$$G^k(a,b) = R^k(abba). \qquad (4.23)$$

Using the symmetry rules of the 3-j symbols,

$$\begin{aligned}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} &= (-)^J \begin{pmatrix} j_1 & j_3 & j_2 \\ m_1 & m_3 & m_2 \end{pmatrix} \qquad \Big| J = j_1 + j_2 + j_3 \\ &= (-)^J \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \end{aligned} \qquad (4.24)$$

and noting that q is fixed by $q = m - m'$,

$$\begin{aligned} d^k(j'm';jm) &= (-)^{m+\frac{1}{2}}\left[(2j+1)(2j'+1)\right]^{1/2} \begin{pmatrix} j & k & j' \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix} \\ &= (-)^{m+\frac{1}{2}}\left[(2j+1)(2j'+1)\right]^{1/2} \begin{pmatrix} j & k & j' \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} j' & k & j \\ m' & q & -m \end{pmatrix} \\ &= (-)^{m+\frac{1}{2}}\left[(2j+1)(2j'+1)\right]^{1/2} \begin{pmatrix} j & k & j' \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} j' & k & j \\ -m' & q & m \end{pmatrix} \\ &= (-)^{-m+m'} d^k(jm;j'm') \end{aligned} \tag{4.25}$$

$\Rightarrow$

$$d^k(j_a m_a; j_b m_b) = (-)^q d^k(j_b m_b; j_a m_a). \tag{4.26}$$

For one electron outside a filled shell, j,

$$\sum_m a^k(jm;j'm') = (2j+1)\delta(k,0). \tag{4.27}$$

Thus the $k > 0$ Coulomb terms vanish in case the potential is spherical. The exchange in this case,

$$\sum_m b^k(jm;j'm') = \frac{1}{2}(2j+1)\Gamma_{jkj'}, \tag{4.28}$$

with

$$\Gamma_{jkj'} = 2\begin{pmatrix} j & k & j' \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}^2. \tag{4.29}$$

The total energy for a closed-shell atom becomes

$$E = E^0 + E^C, \tag{4.30}$$

$$E^0 = \sum_A q_A I(A), \quad q_A = 2j_A + 1, \tag{4.31}$$

$$\begin{aligned} E^C = \sum_A \Big\{ & \underbrace{\frac{1}{2} q_A (q_A - 1) F^0(A,A)}_{\text{a.}} \\ & - \frac{1}{2} q_A \underbrace{\left(\frac{q_A - 1}{2j_A}\right)}_{\text{b.}} \sum_{k>0} \underbrace{(2j_A+1)\frac{1}{2}\Gamma_{j_A k j_A} F^k(A,A)}_{\text{c.}} \Big\} \\ & + \frac{1}{2} \sum_{A,B\ A\neq B} q_A q_B \Big\{ \underbrace{F^0(A,B)}_{\text{d.}} - \sum_k \underbrace{\frac{1}{2}\Gamma_{j_A k j_B} G^k(A,B)}_{\text{e.}} \Big\} \end{aligned} \tag{4.32}$$

a. $k = 0$ part of the intrashell repulsion

b. 1 for a full shell, 0 for one electron

c. Exchange of one j_A electron with the filled shell A

d. $k = 0$ intershell part (the only Coulomb term)

e. The exchange of one j_A electron with the average shell B electron

Example 11 The intrashell interactions of one $p_{3/2}$ shell:
All radial integrals are identical. The angular parts $(m\, m')$ become

$$\begin{aligned}\sum a = a^k(\tfrac{3}{2}\,\tfrac{1}{2}) + a^k(\tfrac{3}{2}\,-\tfrac{1}{2}) + a^k(\tfrac{3}{2}\,-\tfrac{3}{2}) \\ + a^k(\tfrac{1}{2}\,-\tfrac{1}{2}) + a^k(\tfrac{1}{2}\,-\tfrac{3}{2}) \\ + a^k(-\tfrac{1}{2}\,-\tfrac{3}{2})\end{aligned} \tag{4.33}$$

$$\begin{aligned}\sum b = b^k(\tfrac{3}{2}\,\tfrac{1}{2}) + b^k(\tfrac{3}{2}\,-\tfrac{1}{2}) + b^k(\tfrac{3}{2}\,-\tfrac{3}{2}) \\ + b^k(\tfrac{1}{2}\,-\tfrac{1}{2}) + b^k(\tfrac{1}{2}\,-\tfrac{3}{2}) \\ + b^k(-\tfrac{1}{2}\,-\tfrac{3}{2})\end{aligned} \tag{4.34}$$

Note that there are no self-interacting terms, $a^k(\frac{3}{2}\,\frac{3}{2})$ etc.

<u>Case a</u>: $k = 0$

$k = 0 \Rightarrow q = 0 \Rightarrow m = m'$. No self-interaction $\Rightarrow$ the exchange term $b^0(jm, jm') = 0$.

There are six $a^0(m, m')$ in (4.34). Their value equals 1. See I.P. Grant, *Proc. Roy. Soc. Lond. A* **262** (1961) 555, Tables 1-3 or use (4.19)+(4.25).

E.g.

$$
\begin{aligned}
a^0(\tfrac{3}{2}\,\tfrac{1}{2}) &= d^0(\tfrac{3}{2}\,\tfrac{3}{2},\tfrac{3}{2}\,\tfrac{3}{2})d^0(\tfrac{3}{2}\,\tfrac{1}{2},\tfrac{3}{2}\,\tfrac{1}{2}) \\
&= (-)^{\frac{3}{2}+\frac{1}{2}}(2\frac{3}{2}+1)\begin{pmatrix} \frac{3}{2} & 0 & \frac{3}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{3}{2} & 0 & \frac{3}{2} \\ -\frac{3}{2} & 0 & \frac{3}{2} \end{pmatrix} \times \\
&\quad (-)^{\frac{1}{2}+\frac{1}{2}}(2\frac{3}{2}+1)\begin{pmatrix} \frac{3}{2} & 0 & \frac{3}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{3}{2} & 0 & \frac{3}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \qquad (4.35) \\
&= -16\begin{pmatrix} \frac{3}{2} & \frac{3}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}^2\begin{pmatrix} \frac{3}{2} & \frac{3}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} \frac{3}{2} & \frac{3}{2} & 0 \\ \frac{3}{2} & -\frac{3}{2} & 0 \end{pmatrix} \\
&= -16\cdot(-\frac{1}{2})^2(-\frac{1}{2})\frac{1}{2} \qquad \Big|\text{See KK I, p. 109} \\
&= 1 \quad \square
\end{aligned}
$$

(Grant, *op.cit.*, Table 2, same result). $6 \times 1 = 6$. The first term of E^C, eq. (4.32), gives

$$\frac{1}{2}q(q-1) = \frac{1}{2}\cdot 4(4-1) = 6 \Rightarrow 6F^0(A,A). \qquad (4.36)$$

<u>Case b</u>: $k = 2$

Similarly,

$$
\begin{aligned}
&\sum a^2 = \\
&a^2(\tfrac{3}{2}\,\tfrac{1}{2}) + a^2(\tfrac{3}{2}\,-\tfrac{1}{2}) + a^2(\tfrac{3}{2}\,-\tfrac{3}{2}) + a^2(\tfrac{1}{2}\,-\tfrac{1}{2}) + a^2(\tfrac{1}{2}\,-\tfrac{3}{2}) + a^2(-\tfrac{1}{2}\,-\tfrac{3}{2}) \\
&= -\frac{1}{25} \qquad -\frac{1}{25} \qquad +\frac{1}{25} \qquad +\frac{1}{25} \qquad -\frac{1}{25} \qquad -\frac{1}{25} \\
&= -\frac{2}{25} \qquad (4.37)
\end{aligned}
$$

Symmetries:

$$
\begin{aligned}
a^k(jm;j'm') &= a^k(j'm';jm) \quad &= a^k(j-m;j'-m) \\
&= a^k(j-m;j'm') &= a^k(jm;j'-m') \qquad (4.38) \\
b^k(jm;j'm') &= b^k(j'm';jm) \quad &= b^k(j-m;j'-m') \qquad (4.39)
\end{aligned}
$$

The coefficients b^k are given in Table 4 of Grant. Here,

$$\begin{aligned}
&\sum b^2 = \\
&b^2(\tfrac{3}{2}\,\tfrac{1}{2}) + b^2(\tfrac{3}{2}\,-\tfrac{1}{2}) + b^2(\tfrac{3}{2}\,-\tfrac{3}{2}) + b^2(\tfrac{1}{2}\,-\tfrac{1}{2}) + b^2(\tfrac{1}{2}\,-\tfrac{3}{2}) + b^2(-\tfrac{1}{2}\,-\tfrac{3}{2}) \\
&= \frac{2}{25} \qquad + \frac{2}{25} \qquad + 0 \qquad + 0 \qquad + \frac{2}{25} \qquad + \frac{2}{25} \\
&= \frac{8}{25}
\end{aligned} \tag{4.40}$$

The radial parts $R(abab) = R(abba) = R(aaaa)$. Adding,

$$\sum a^2 - \sum b^2 = (-\frac{2}{25}) - (\frac{8}{25}) = -\frac{2}{5} \tag{4.41}$$

In eq. (4.32),

$$-\frac{1}{2}q(\frac{q-1}{2j})(2j+1)\frac{1}{2}\Gamma_{\frac{3}{2}2\frac{3}{2}} = -\frac{1}{2}\cdot 4\cdot\frac{4-1}{3}\cdot 4\cdot\frac{1}{2}\cdot\underbrace{\frac{1}{10}}_{\text{Grant, Table 5, p. 570}} = -\frac{2}{5} \tag{4.42}$$

4.2 The Dirac-Fock equations

The two-electron energy (4.32) equals

$$\begin{aligned}
E^C = &\frac{1}{2}\sum_A q_A \int_0^\infty \frac{\mathrm{d}r}{r}\Big[P_A^2(r) + Q_A^2(r)\Big]\times \\
&\Bigg\{(q_A-1)r\int_0^\infty \frac{1}{r_>}\mathrm{d}s\Big[P_A^2(s)+Q_A^2(s)\Big] \qquad \Big|\text{a.} \\
&+\sum_{B\neq A} q_B r\int_0^\infty \frac{1}{r_>}\mathrm{d}s\Big[P_B^2(s)+Q_B^2(s)\Big] \qquad \Big|\text{b.} \\
&-\sum_{k>0}\frac{q_A-1}{2j_A}\frac{1}{2}(2j_A+1)\Gamma_{j_A k j_A} r\int_0^\infty \frac{r_<^k}{r_>^{k+1}}\Big[P_A^2(s)+Q_A^2(s)\Big]\mathrm{d}s\Bigg\} \Big|\text{c.} \\
&-\frac{1}{2}\sum_{A\neq B} q_A q_B \sum_k \frac{1}{2}\Gamma_{j_A k j_B}\int_0^\infty \mathrm{d}r\int_0^\infty \mathrm{d}s\frac{r_<^k}{r_>^{k+1}} \qquad \Big|\text{d.} \\
&\quad\times\Big[P_A(s)P_B(s)+Q_A(s)Q_B(s)\Big]\Big[P_A(r)P_B(r)+Q_A(r)Q_B(r)\Big].
\end{aligned} \tag{4.43}$$

When introducing the variation, ΔQ, a factor of 4 will appear in each of the four terms, from:

a. $\Delta(Q^2) = 2Q\Delta Q$, both $\Delta Q(r)$ and $\Delta Q(s)$

b. $\Delta(Q^2) = 2Q\Delta Q$, $A = \mathrm{i}$ or $B = \mathrm{i}$

c. As **a.**

d. $Q(r)$ or $Q(s)$, $A = \mathrm{i}$ or $B = \mathrm{i}$

(4.44)

In terms of the variations ΔP and ΔQ, the variation of the two-electron energy becomes

$$\begin{aligned} E^C =& 2q_A \int_0^\infty \frac{Dr}{r}\Big[P_A\Delta P_A + Q_A\Delta Q_a\Big]\times \\ &\Bigg\{(q_A-1)r\int_0^\infty \frac{1}{r_>}\mathrm{d}s\Big[P_A^2(s)+Q_A^2(s)\Big] \\ &+\sum_{B\neq A} q_B r\int_0^\infty \frac{1}{r_>}\mathrm{d}s\Big[P_B^2(s)+Q_B^2(s)\Big] \\ &-\sum_{k>0}\frac{q_A-1}{2j_A}\frac{1}{2}(2j_A+1)\Gamma_{j_Akj_A}r\int_0^\infty \frac{r_<^k}{r_>^{k+1}}\Big[P_A^2(s)+Q_A^2(s)\Big]\mathrm{d}s\Bigg\} \\ &-\sum_{B\neq A} q_Aq_B\sum_k \Gamma_{j_Akj_B}\int_0^\infty \mathrm{d}r\int_0^\infty \mathrm{d}s \\ &\times\Big[P_A(s)P_B(s)+Q_A(s)Q_B(s)\Big]\Big[P_B(r)P\Delta P_A + Q_B(r)\Delta Q_A\Big]. \end{aligned} \tag{4.45}$$

We introduce the auxiliary function

$$\begin{aligned} Y_n(A,r) =& -(q_A-1)Y_0(A,A;r) - \sum_{B\neq A} q_BY_0(B,B;r) \\ &+\sum_{k>0}\frac{q_A-1}{2j_A}\frac{1}{2}(2j_A+1)\Gamma_{j_Akj_A}Y_k(A,A;r), \end{aligned} \tag{4.46}$$

where

$$Y_k(ac;r) = r\int_0^\infty \frac{r_<^k}{r_>^{k+1}}\Big[P_a(s)P_c(s)+Q_a(s)Q_c(s)\Big]\mathrm{d}s \tag{4.13}$$

could be interpreted as an *effective electronic charge*, containing both inner and outer screening.

Orthogonality leads to the Lagrange coefficients ε, whose variation

$$\Delta\varepsilon = \Delta\Big[E - \sum_A q_A\varepsilon_{AA}N_{AA} - \sum_{A\neq B} q_Aq_B\varepsilon_{AB}N_{AB}\Big], \tag{4.47}$$

where

$$N_{AB} = \int_o^\infty (P_AP_B + Q_AQ_B)\mathrm{d}r. \tag{4.48}$$

Introducing $\varepsilon_A = mc^2 - E_A$, eq. (3.134) gives

$$\begin{cases} c(P' + \frac{\kappa}{r}P) + (-mc^2 + V - mc^2 - \varepsilon)Q = 0 & \Big| + Q\times \\ c(Q' - \frac{\kappa}{r}Q) + (-mc^2 - V + mc^2 + \varepsilon)P = 0 & \Big| - P\times \end{cases} \tag{4.49}$$

$$\begin{aligned} \varepsilon(P^2 + Q^2) &= c\Big[QP' + \frac{\kappa}{r}QP - PQ' + \frac{\kappa}{r}PQ\Big] - 2mc^2Q^2 \\ &\quad + V[P^2 + Q^2] \\ &= c\Big[QP' - PQ' + \frac{2\kappa}{r}PQ\Big] - 2mc^2Q^2 + V[P^2 + Q^2]. \end{aligned} \tag{4.50}$$

The variation of this one-electron term,

$$\begin{aligned} \Delta\varepsilon_A = q_A \int_0^\infty \Big\{ & c\Big[\Delta QP' + Q\Delta P' - \Delta PQ' - P\Delta Q' + \frac{2\kappa}{r}(P\Delta Q + Q\Delta P)\Big] \\ & -4c^2Q\Delta Q + 2V(P\Delta P + Q\Delta Q)\Big\}\mathrm{d}r. \end{aligned} \tag{4.51}$$

From $\Delta P = \Delta Q = 0$ we get by partial integration

$$\begin{cases} \int_0^\infty Q\Delta P'\mathrm{d}r = -\int_0^\infty Q'\Delta P\mathrm{d}r \\ \int_0^\infty P\Delta Q'\mathrm{d}r = -\int_0^\infty P'\Delta Q\mathrm{d}r \end{cases} \tag{4.52}$$

Hence the variation of the one-electron energy

$$\begin{aligned} \Delta\varepsilon_A = 2q_Ac \int_0^\infty \Big\{ & \Delta Q_AP'_A + \Delta P'_AQ'_A + \frac{\kappa}{r}(P_A\Delta Q_A + Q_A\Delta P_A) \\ & -2cQ_A\Delta Q_A + \frac{1}{c}V(P_A\Delta P_A + Q_A\Delta Q_A)\Big\}\mathrm{d}r. \end{aligned} \tag{4.53}$$

Adding to this the variation of the two-electron energy, ΔE^C, eq. (4.45), and subtracting the Lagrange multiplier terms for normalization and orthogonality, we get

$$\int_0^\infty 2q_Ac\Big\{\Delta Q_AP'_A-\Delta P_AQ'_A+\frac{\kappa_A}{r}(P_A\Delta Q_A+Q_A\Delta P_A) \tag{4.54}$$

$$-2cQ_A\Delta Q_A+\frac{1}{c}V_n(P_A\Delta P_A+Q_A\Delta Q_A)\Big\} \quad \Big|\text{a.}$$

$$-2q_A\int_0^\infty\frac{\mathrm{d}r}{r}\Big[P_A\Delta P_A+Q_A\Delta Q_A\Big]Y_n(A,r) \quad \Big|\text{b.}$$

$$-\sum_{B\neq A}q_Aq_B\sum_k\Gamma_{j_Akj_B}\int_0^\infty\frac{\mathrm{d}r}{r}Y_k(A,B;r)[P_B\Delta P_A+Q_B\Delta Q_A] \Big|\text{c.}$$

$$-2q_A\varepsilon_{AA}\int_0^\infty(P_A\Delta P_A+Q_A\Delta Q_A)\mathrm{d}r \quad \Big|\text{d.}$$

$$-2\sum_{B\neq A}q_Aq_B\varepsilon_{AB}\int_0^\infty(\Delta P_AP_B+\Delta Q_AQ_B)\mathrm{d}r=0. \quad \Big|\text{e.}$$

a. one-electron

b. use (4.46)

c. (4.45), inter-shell exchange

d. normalization

e. orthogonality

In the last term (**e.**), a factor of 2 comes from the possibilities "A" $= A$ and "B" $= A$. Dividing (4.54) by $2q_Ac$ the coefficient of ΔQ_A and $-\Delta P_A$ become

$$\boxed{\begin{aligned}&P'_A+\frac{\kappa_A}{r}P_A+\frac{1}{c}[-2c^2+V_n-\frac{1}{r}Y_n(A,r)-\varepsilon_{AA}]Q_A\\&+\sum_{B\neq A}q_B\Big[-\sum_k\Gamma_{j_Akj_B}\frac{1}{2cr}Y_k(A,B;r)-\frac{1}{c}\varepsilon_{AB}\delta(\kappa_A,\kappa_B)\Big]Q_B=0\\&Q'_A+\frac{\kappa_A}{r}Q_A+\frac{1}{c}[-V_n+\frac{1}{r}Y_n(A,r)+\varepsilon_{AA}]P_A\\&+\sum_{B\neq A}q_B\Big[\sum_k\Gamma_{j_Akj_B}\frac{1}{2cr}Y_k(A,B;r)+\frac{1}{c}\varepsilon_{AB}\delta(\kappa_A,\kappa_B)\Big]P_B=0\end{aligned}} \tag{4.55}$$

These are the Dirac-Fock equations for a closed-shell atom. We may set

$$\frac{1}{r}Y(A,r)=-V_n+\frac{1}{r}Y_n(A,r)=\frac{1}{r}[Z+Y_n(A,r)] \tag{4.56}$$

4.2.1 The relativistic Koopmans theorem

Multiply (4.55a) by cQ_A, (4.55b) by $-cP_A$, add, integrate and suppose $\varepsilon_{AB} = 0$:

$$\int_0^\infty \Bigg[cQ_A\Big(P'_A + \frac{\kappa}{r}P_A\Big) - cP_A\Big(Q'_A - \frac{\kappa_A}{r}Q_A\Big) - 2c^2Q_A^2 \qquad (4.57)$$
$$+ V_n(P_A^2 + Q_A^2) - \varepsilon_{AA}(P_A^2 + Q_A^2) + (P_A^2 + Q_A^2)\Big(-\frac{1}{r}Y_n(A,r)\Big)$$
$$+ c\sum_{B\neq A} q_B(Q_AQ_B + P_AP_B)\sum_k \Gamma_{j_Akj_B}\frac{1}{2cr}Y_k(A,B;r)\Bigg]\mathrm{d}r = 0.$$

$\Rightarrow$

$$\begin{aligned}
\varepsilon_{AA} = \; & \int_0^\infty \Big[cQ_A\Big(P'_A + \frac{\kappa}{r}P_A\Big) - cP_A\Big(Q'_A - \frac{\kappa}{r}Q_A\Big) - 2c^2Q_A^2 + V_n\Big] \;\Big| I_A \\
& + \int_0^\infty \mathrm{d}r(P_A^2 + Q_A^2)\Bigg[+\frac{q_A - 1}{r}Y_0(A,A;r) + \sum_{B\neq A}\frac{1}{r}q_BY_0(B,B;r)\Big| [4.46] \\
& \qquad\qquad - \sum_{k>0}\frac{q_A - 1}{2j_Ar}\frac{1}{2}(2j_A+1)\Gamma_{j_Akj_A}Y_k(A,A;r)\Bigg] \\
& + c\sum_{B\neq A} q_B(Q_A(r)Q_B(r) + P_A(r)P_B(r))\sum_k \Gamma_{j_Akj_B}\frac{1}{2cr}\times \\
& \qquad\qquad r\int_0^\infty \frac{r_<^k}{r_>^{k+1}}[P_A(s)P_B(s) + Q_A(s)Q_B(s)]\mathrm{d}s \qquad\Big| [4.13] \\
= \; & I_A + (q_A - 1)F^0(A,A) + \sum_{B\neq A} q_BF^0(A,B) \\
& - \sum_{k>0}\frac{q_A - 1}{2j_A}\frac{1}{2}(2j_A+1)\Gamma_{j_Akj_A}F^k(A,A) \\
& - \sum_{B\neq A}\frac{q_B}{2}\sum_k \Gamma_{j_Akj_B}G^k(A,B). \qquad (4.58)
\end{aligned}$$

A comparison with the total energy, (4.32) shows this to be the difference $E(q_A) - E(q_A - 1)$ $\square$

4.2.2 Multiconfiguration treatment, a simple example

(C. Briançon, J.P. Desclaux, *Phys. Rev. A* **15** (1976) 2157)

The total energy for a $p^4(J=2)$ configuration. At the LS-limit,

$$\begin{cases} |^3\mathrm{P}_2\rangle = \frac{1}{\sqrt{3}}\Big[\sqrt{2}|\bar{p}^2p^2\rangle + \quad |\bar{p}p^3\rangle\Big] \\ |^2\mathrm{D}_2\rangle = \frac{1}{\sqrt{3}}\Big[\quad |\bar{p}^2p^2\rangle - \sqrt{2}|\bar{p}p^3\rangle\Big] \end{cases} \tag{4.59}$$

Generally,

$$|J=2\rangle = a|\bar{p}^2p^2\rangle + \sqrt{1-a^2}|\bar{p}p^3\rangle, \tag{4.60}$$

with the total energy

$$\begin{aligned} E_T = {} & a^2E_{\mathrm{av}}(\bar{p}^2p^2) + (1-a^2)E_{\mathrm{av}}(\bar{p}p^3) \\ & -\frac{4}{75}a^2F^2(p,p) + \frac{3}{50}(1-a^2)G^2(\bar{p},p) \\ & +\frac{8}{50}\sqrt{2}a\sqrt{1-a^2}R^2(\bar{p},\bar{p},p,p). \end{aligned} \tag{4.61}$$

4.2.3 "Average-of-configuration" treatment

Assume that all $|jm\rangle$ states of the one-electron states are equally probable in the many-electron wave function.

Example 21 Iodine, $5p^5$.

The hole can be in the $5p$ or the $5\bar{p}$ shell. They have 4 and 2 m_j-states, respectively. Therefore the statistical weights are $\frac{4}{6}$ and $\frac{2}{6}$, respectively.

Example 32 Ground state of uranium, $5f^36d$.

a) f-configuration.

$$\begin{aligned} l=3, \quad f:& j=3+\frac{1}{2}=\frac{7}{2}, \quad 2j+1=8 \\ \bar{f}:& j=3-\frac{1}{2}=\frac{5}{2}, \quad 2j+1=6 \end{aligned}$$

config	$\bar{f}^3$	$\bar{f}^2f$	$\bar{f}f^2$	f^3	$\sum$
weight	$\frac{6\cdot5\cdot4}{3!}=20$	$\frac{6\cdot5}{2!}\cdot\frac{8}{1!}=120$	$\frac{6}{1!}\cdot\frac{8\cdot7}{2!}=168$	$\frac{8\cdot7\cdot6}{3!}=56$	364
	$=\frac{5}{91}$	$=\frac{30}{91}$	$=\frac{42}{91}$	$=\frac{14}{91}$	

Note that you have to include the permutation of electrons, $n!$, in the weight calculation.

b) d-configuration

$$\begin{aligned} l=2, \quad d:& j=2+\frac{1}{2}=\frac{5}{2}, \quad 2j+1=6 \\ \bar{d}:& j=2-\frac{1}{2}=\frac{3}{2}, \quad 2j+1=4 \end{aligned}$$

config	$\bar{d}$	d	$\sum$
weight	4	6	10
	$= \frac{2}{5}$	$= \frac{3}{5}$	

Altogether eight jj-coupled configurations:

config	$\bar{d}\bar{f}^3$	$\bar{d}\bar{f}^2 f$	$\bar{d}\bar{f}f^2$	$\bar{d}f^3$
weight	$\frac{2}{5}\cdot\frac{5}{91}$	$\frac{2}{5}\cdot\frac{30}{91}$	$\frac{2}{5}\cdot\frac{42}{91}$	$\frac{2}{5}\cdot\frac{14}{91}$
weight$\times$455	10	60	84	28
config	$d\bar{f}^3$	$d\bar{f}^2 f$	$d\bar{f}f^2$	df^3
weight	$\frac{3}{5}\cdot\frac{5}{91}$	$\frac{3}{5}\cdot\frac{30}{91}$	$\frac{3}{5}\cdot\frac{42}{91}$	$\frac{3}{5}\cdot\frac{14}{91}$
weight$\times$455	15	90	126	42

4.3 Numerical solution of the Dirac-Fock equations

Choosing the variable

$$t = \ln r \tag{4.62}$$

and the step length $h = \Delta t = 0.05$, one usually obtains a precision of about 9 figures (for E_T). About 200 radial grid points, r_i, are needed between the first point and the practical infinity. All wave functions P_A, Q_A and potentials $\frac{1}{r}Y(A,r)$ and $\frac{1}{r}\sum_{B\neq A} q_B \sum_k \Gamma_{j_A k j_B} Y_k(A,B;R)$ are expressed at the r_i.

The integrations outwards are started by the series expansion

$$\begin{cases} P(r) = r^\gamma(p_0 + p_1 r + p_2 r^2 + \quad \ldots) \\ Q(r) = r^\gamma(q_0 + q_1 r + q_2 r^2 + \quad \ldots). \end{cases} \tag{4.63}$$

After sufficiently many points are obtained, a five-point **Adams predictor-corrector method** or some other numerical method is used. Consider

$$y' = \frac{\mathrm{d}y}{\mathrm{d}t} = f(y)y(t) + g(t) \tag{4.64}$$

The new, $(n+1)$:th point is obtained by combining the *predictor*

$$p_{n+1} = y_n + \frac{h}{720}\left[1901y'_n - 2774y'_{n-1} + 2616y'_{n-2} - 1274y'_{n-3} + 251y'_{n-4}\right] \tag{4.65}$$

and the *corrector*

$$c_n + 1 = y_n + \frac{h}{720}\left[251p'_{n+1} + 646y'_n - 264y'_{n-1} + 106y'_{n-2} - 19y'_{n-3}\right] \tag{4.66}$$

to

$$y_{n+1} = \frac{1}{502}\left[473c_{n+1} + 29p_{n+1}\right] + \mathcal{O}(h^7) \tag{4.67}$$

The outwards integration becomes unstable after the (effective, radial) potential, $V(r) = -\frac{1}{r}Y(A,r)$ crosses the eigenvalue, ε

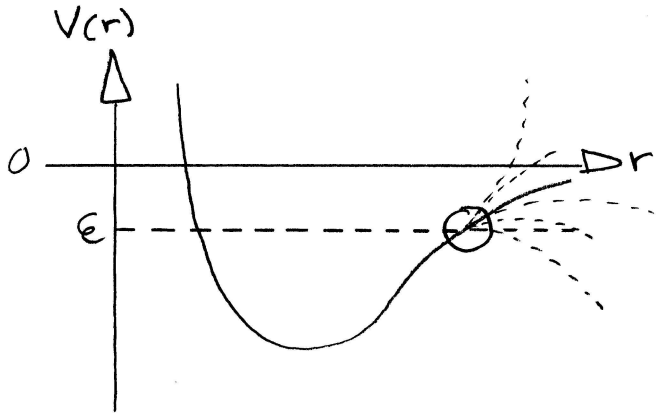

Figure 4.1: The instability of the integration depicted

Similarly, the inwards integration becomes unstable after this point. The solution is to fit the two integrations at this point.

In "Method 2" one solves both the original inhomogeneous differential equations and the homogeneous ones with $\sum_{B\neq A} = 0$. Then every linear combination of the two solutions is a solution.

The coefficients of the homogeneous solution, α, follow from

$$\begin{cases} P^{\mathrm{i}}_{\mathrm{L}}(r_{\mathrm{I}}) + \alpha_{\mathrm{L}} P^{\mathrm{h}}_{\mathrm{L}}(r_{\mathrm{I}}) = P^{\mathrm{i}}_{\mathrm{R}}(r_{\mathrm{I}}) + \alpha_{\mathrm{R}} P^{\mathrm{h}}_{\mathrm{R}}(r_{\mathrm{I}}) \\ Q^{\mathrm{i}}_{\mathrm{L}}(r_{\mathrm{I}}) + \alpha_{\mathrm{L}} Q^{\mathrm{h}}_{\mathrm{L}}(r_{\mathrm{I}}) = Q^{\mathrm{i}}_{\mathrm{R}}(r_{\mathrm{I}}) + \alpha_{\mathrm{R}} Q^{\mathrm{h}}_{\mathrm{R}}(r_{\mathrm{I}}), \end{cases} \tag{4.68}$$

at the matching point, r_{I}. Here i=inhomogeneous, h=homogeneous, L="left" ($r < r_{\mathrm{I}}$), R="right" ($r > r_{\mathrm{I}}$).

This solution fulfills the boundary conditions but may not be normalized. Normalization by a constant is not allowed because the equation is inhomogeneous $\Rightarrow$ the eigenvalue of ε_{AA} is chosen to obtain normalization.

Actually it suffices if we converge towards a normalized solution. Let $\varepsilon = \varepsilon_{AA}$. Then

$$\int_0^\infty (P^2 + Q^2)\mathrm{d}r + 2\delta\varepsilon \int_0^\infty \Big(P\frac{\partial P}{\partial \varepsilon} + Q\frac{\partial Q}{\partial \varepsilon}\Big)\mathrm{d}r = 1. \tag{4.69}$$

By first differentiating the original Dirac-Fock equations,

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial P}{\partial \varepsilon}\Big) + \kappa\dfrac{\partial P}{\partial \varepsilon} + f(r)\dfrac{\partial Q}{\partial \varepsilon} = \quad \dfrac{1}{2c}Q \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial Q}{\partial \varepsilon}\Big) - \kappa\dfrac{\partial Q}{\partial \varepsilon} + g(r)\dfrac{\partial P}{\partial \varepsilon} = -\dfrac{1}{2c}P, \end{cases} \tag{4.70}$$

these derivatives can be solved by the same subroutines as P and Q themselves. Then (4.69) $\Rightarrow$

$$\delta\varepsilon_A = \frac{1 - \int_0^\infty (P_A^2 + Q_A^2)\mathrm{d}r}{2\int_0^\infty \Big(P_A\frac{\partial P_A}{\partial \varepsilon_A} + Q_A\frac{\partial Q_A}{\partial \varepsilon_A}\Big)\mathrm{d}r}. \tag{4.71}$$

In the old Dirac-Fock One-Center Expansion (DF-OCE) method, four numerical methods could be used:

1.
- Calculate only $\psi_{\mathrm{R}}^{\mathrm{h}}$, put $P_{\mathrm{L}} = P_{\mathrm{R}}^{\mathrm{i}} + \alpha_{\mathrm{R}} P_{\mathrm{R}}^{\mathrm{h}}$.
- Fit Q by changing ε.
- Finally normalize by force (inner shells) or by using $\partial P/\partial\varepsilon$ (molecules).

2. Previous method. Normalize by adding to P some amount of $\delta\varepsilon \times (\partial P/\partial\varepsilon)$. If the number of nodes requires it, change ε. The continuity is guaranteed.

- Fit ε by (4.71)

3. Use both α_{L} and α_{R}. Use (4.71). For wrong number of nodes,

4. change ε to get the right number of nodes.

4.3.1 Specific features of the DF-OCE method

- Each MO consists of a *dominant* AO and other AOs. **Example**: For the Γ_7 MO of CH_4, $2\bar{p}_{3/2} + 4d + 4f$.
- Use for the dominant component Method 1 and for the others Method 3 or, if necessary, Method 4.
- Then the dominant component fixes ε (which is not changed in Method 3 for the other components).

- After treating the entire MO, set the common

$$\varepsilon_{\mathrm{MO}} = \sum_{\mathrm{AO}} N(\mathrm{AO})\varepsilon_{\mathrm{AO}} \tag{4.72}$$

(SOLDMO, lines 175-177)

Then correct it by setting

$$\varepsilon_{\mathrm{MO}} \to \varepsilon_{\mathrm{MO}} + \frac{1 - \sum_i N_i}{\sum_i 2\int_0^\infty (P_i p_i + Q_i q_i)} \Bigg| p_i = \frac{\partial P_i}{\partial \varepsilon_i} \tag{4.73}$$

(SOLDMO, lines 178, 180)

- After treating the entire MO, normalize it by the coefficient $[\sum_i N_i]^{-1/2}$. (SOLDMO, lines 157, 204-211)

Some results:

Desclaux and Pyykkö (1974):	the bond-length contraction
D&P (1976):	the Ag/Au difference
D&P (1976):	$\bar{p}$ bonding in TlH.
D&P (1977a):	TiH_4–$(104)H_4$. Trends.
D&P (1977b):	$(114)H_4 \gtrsim PbH_4$.
D&P (1978):	MoH_6/WH_6. R_e, D_e, k_2; oxidation state.
P (1979b):	MH^+, MH_2, Groups 2 and 12.
P (1979c):	LnH, AnH.

[See Table 7.3 of RTAM I.]

The iterations can be speeded up (by roughly a factor of 2) by using a *convergence factor* α.

Let $P_n^{\mathrm{i}}, Q_n^{\mathrm{i}}$ and $P_n^{\mathrm{f}}, Q_n^{\mathrm{f}}$ be the initial and final values during the n:th iteration. Choose

$$\begin{cases} P_{n+1}^{\mathrm{i}}(r) = \beta[\alpha P_n^{\mathrm{f}} + (1-\alpha)P_n^{\mathrm{i}}] \\ Q_{n+1}^{\mathrm{i}}(r) = \beta[\alpha Q_n^{\mathrm{f}} + (1-\alpha)Q_n^{\mathrm{i}}] \end{cases} \tag{4.74}$$

Here β is a normalization factor and α is restricted to

$$0.1 \le \alpha \le 0.9$$

and is calculated every time from the change

$$\Delta R_n = \max[P_n^{\mathrm{i}}(r) - P_n^{\mathrm{f}}(r), Q_n^{\mathrm{i}}(r) - Q_n^{\mathrm{f}}(r)].$$

5 Symmetry

5.1 Rotation Operators

Consider a scalar field $f(\underline{r})$.

Example 1: The temperature at point $\underline{r}$

After a rotation of the coordinate system, the same field is described by the new function $f'(\underline{r}')$, with

$$\underline{r}' = R\underline{r} \tag{5.1}$$

Here the operator R acts on the coordinates. Introduce now the operator P_R, which acts on the function f in Hilbert space, making it follow the rotation of the coordinate system:

$$f'(\underline{r}') = P_R f(\underline{r}') = f(\underline{r}) \tag{5.2}$$

Active interpretation: In the left-hand equation, $f'(\underline{r}') = P_R f(\underline{r}')$, the operator P_R changes the function f into f'.

Passive interpretation: In the right-hand equation, rewritten as $f(\underline{r}') = P_R^{-1} f(\underline{r})$, the operator does not change the function but acts on the coordinates $\underline{r}$, changing them to $\underline{r}'$

Example 2: Let $f(\phi) = \mathrm{e}^{\mathrm{i}m\phi}$ and $R\phi = \phi' = \phi - \alpha$, a positive rotation through α:

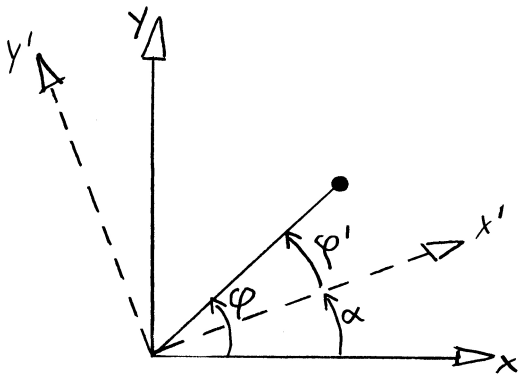

Figure 5.1: Positive rotation through α

Then

$$\underbrace{\mathrm{e}^{\mathrm{i}m(\phi'+\alpha)}}_{f'(\phi')} = \underbrace{\mathrm{e}^{\mathrm{i}m\alpha}}_{P_R}\underbrace{\mathrm{e}^{\mathrm{i}m\phi'}}_{f(\phi')} = \underbrace{\mathrm{e}^{\mathrm{i}m\phi}}_{f(\phi)} \tag{5.3}$$

5.1.1 The Euler Angles

Consider a right-handed coordinate system $\mathbf{O}xyz$ and positive rotations (like in fig. (5.1), R_z pushing a right-handed screw towards $+z$). Introduce the rotations

(i) Rotation $R_z(\alpha)$ through α around $\mathbf{O}_z$ $(\mathbf{O}_y \to \mathbf{O}_{y'})$
(ii) Rotation $R_{y'}(\beta)$ through β around $\mathbf{O}_{y'}$ $(\mathbf{O}_z \to \mathbf{O}_{z''})$
(iii) Rotation $R_{z''}(\gamma)$ through γ around $\mathbf{O}_{z''}(\mathbf{O}_{y'} = \mathbf{O}_{y''} \to \mathbf{O}_{y'''})$

Then eq. (5.3) $\Rightarrow$ (replacing m by the operator j_z)

$$f'(\underline{r}') = \mathrm{e}^{\mathrm{i}\alpha j_{z'}} f(\underline{r}') \tag{5.4}$$

$$\begin{aligned} f''(\underline{r}'') &= \mathrm{e}^{\mathrm{i}\beta j_{y''}} f'(\underline{r}') \\ &= \mathrm{e}^{\mathrm{i}\beta j_{y''}} \mathrm{e}^{\mathrm{i}\alpha j_{z''}} f(\underline{r}'') \end{aligned} \tag{5.5}$$

$$\begin{aligned} f'''(\underline{r}''') &= \mathrm{e}^{\mathrm{i}\gamma j_{z'''}} f''(\underline{r}'') \\ &= \mathrm{e}^{\mathrm{i}\gamma j_{z'''}} \mathrm{e}^{\mathrm{i}\beta j_{y'''}} \mathrm{e}^{\mathrm{i}\alpha j_{z'''}} f''(\underline{r}''') \end{aligned} \tag{5.6}$$

in the *active interpretation* using *temporary coordinates* $\mathbf{O}_z$, $\mathbf{O}'_y$.

The P_R operator of the type (5.2)-(5.3) becomes

$$P_R(\alpha\beta\gamma) = \mathrm{e}^{\mathrm{i}\gamma j_z} \mathrm{e}^{\mathrm{i}\beta j_y} \mathrm{e}^{\mathrm{i}\alpha j_z} \tag{5.7}$$

With this operator, the same coordinates must be used on both sides.

If fixed rotation axes, $\mathbf{O}xyz$, are used, one should replace $(\gamma\beta\alpha)$ by $(\alpha\beta\gamma)$ in the operator P_R.

Rotations	Axes	Operations		
		1.(z)	2.(y)	3.(z)
Active	Fixed	γ	β	α
Active	Temporary	α	β	γ
Passive	Fixed	$-\alpha$	$-\beta$	$-\gamma$
Passive	Temporary	$-\gamma$	$-\beta$	$-\alpha$

5.1.2 Rotation of Spherical Harmonics.

In the *active, temporary* picture, the rotated function becomes

$$Z'(\vartheta'\phi') = \mathrm{e}^{-\mathrm{i}\gamma j_{z'}} \mathrm{e}^{\mathrm{i}\beta j_{y'}} \mathrm{e}^{\mathrm{i}\alpha j_{z'}} \mathrm{Y}_{lm}(\vartheta'\phi') \tag{5.8}$$

In the *passive, temporary* representation

$$\begin{aligned} \mathrm{Y}_{lm}(\vartheta'\phi') &= \mathrm{e}^{-\mathrm{i}\alpha j_z}\mathrm{e}^{-\mathrm{i}\beta j_y}\mathrm{e}^{-\mathrm{i}\gamma j_z}\mathrm{Y}_{lm}(\vartheta\phi) \\ &\equiv \sum_{m'} \mathrm{D}^l_{m'm}(\alpha\beta\gamma)\mathrm{Y}_{lm'}(\vartheta\phi) \end{aligned} \tag{5.9}$$

Here the *rotation matrices* (***Wigner*** *matrices*)

$$\mathrm{D}^l_{m'm} = \langle lm'|\mathrm{e}^{-\mathrm{i}\alpha j_z}\mathrm{e}^{-\mathrm{i}\beta j_y}\mathrm{e}^{-\mathrm{i}\gamma j h_z}|lm\rangle \tag{5.10}$$

The temporary rotations 1.α, 2.β, 3.γ bring $\mathbf{O}xyz$ to $\mathbf{O}x'y'z'$.

5.1.3 Rotation of $|jm\rangle$ Functions

Because the $|jm\rangle$ functions, with half-integer j, obey the same commutation rules as those with integer j, we may simply substitute l for j:

$$P_R^{-1}|jm\rangle = \sum_{m'} \mathrm{D}^j_{m'm}(\alpha\beta\gamma)|jm'\rangle \tag{5.11}$$

5.1.4 The 2-to-1 Homomorphism from $SU(2)$ to $SO(3)$

The ***special unitary*** *group,* $SU(2)$, consists of all unitary ($\mathbf{U}^\dagger = \mathbf{U}^{-1}$), unimodular ($\det \mathbf{U} = 1$) 2×2 matrices. We may regard it as the group of all three-dimensional rotations, spanned by the spin vectors

$$\begin{cases} \eta_1 = \alpha = |\frac{1}{2}\ \frac{1}{2}\rangle \\ \eta_2 = \beta = |\frac{1}{2}\ \text{-}\frac{1}{2}\rangle \end{cases} \tag{5.12}$$

A Pauli matrix σ_i transforms under the rotation $\mathbf{U}$ to

$$\mathbf{U}\sigma_i\mathbf{U}^\dagger = \sum_j^{x,y,z} \sigma_j \mathbf{A}_{ji} \tag{5.13}$$

The matrix $\mathbf{A}$ is real, orthogonal ($\mathbf{A}^\dagger\mathbf{A} = 1$) and unimodular and belongs hence to the ***special orthogonal*** *group* $SO(3)$ (consisting of 3×3 matrices), isomorphic to the group $R(3)$ of all three-dimensional rotations.

$$Rm_iR^\dagger = \sum_j m_j \mathbf{A}_{ji} \tag{5.14}$$

where the vectors m_j form a cartesian basis.

Consider now two elements, $\mathbf{U}$ and $\mathbf{V}$, of $SU(2)$ and two elements, $\mathbf{A}$ and $\mathbf{B}$, of $SO(3)$, $\rightarrow R(\mathbf{A})$ and $R(\mathbf{B})$. A comparison of (5.13) and (5.14) yields the mapping

$$\mathbf{U} \rightarrow \mathbf{A} \tag{5.15}$$

Letting $\mathbf{V} \to \mathbf{B}$, (5.13) $\Rightarrow$

$$(\mathbf{UV})\sigma_i(\mathbf{V}^\dagger\mathbf{U}^\dagger) = \mathbf{U}\sigma_j\mathbf{B}_{ji} \tag{5.16}$$

Hence (5.15) is a homomorphism. In particular, as

$$\mathbf{U}\sigma_i\mathbf{U}^\dagger = (-\mathbf{U})\sigma_i(-\mathbf{U})^\dagger = \sigma_j\mathbf{A}_{ji} \Rightarrow \pm\mathbf{U} \to \mathbf{A}, \tag{5.17}$$

it is a 2-to-1 homomorphism, exactly two elements of $SU(2)$ is mapped to exactly one element of $SO(3)$.

As the angular momentum $\underline{j} = \underline{\sigma}/2$ for $SU(2)$, the matrix operator for a rotation through ϕ around $\underline{n}$ becomes (recall (5.3)):

$$\mathbf{U}(\underline{n}, \phi) = \mathrm{e}^{\mathrm{i}(\underline{\sigma}\cdot\underline{n})\phi/2} = \cos\frac{\phi}{2} + \mathrm{i}(\underline{\sigma}\cdot\underline{n})\sin\frac{\phi}{2} \tag{5.18}$$

Hence

$$\mathbf{U}(\underline{n}, 2\pi) = -\underline{\underline{\mathbf{1}}} \tag{5.19}$$

$$\mathbf{U}(\underline{n}, 4\pi) = \underline{\underline{\mathbf{1}}} \tag{5.20}$$

5.2 Double Groups

5.2.1 Non-relativistic Case with Spin

Let G be the non-relativistic or *simple* group. Its elements, like the pure rotations, R_i, commute with $\mathbf{H}$,

$$[\mathbf{H}, R_i] = 0 \tag{5.21}$$

The spin space is entirely decoupled from coordinate space. The total symmetry group is the direct product

$$G_{\text{full}} = G \times SU(2) \tag{5.22}$$

For a Coulomb potential, G_{full} is larger still (**Fock** (1936)).

5.2.2 Relativistic Case

Now the space and spin variables are coupled. Only those $\mathbf{U}_i \in SU(2)$ which correspond to an $\mathbf{R}_i$ (by (5.15)) can be included in

$$G^* = \{\pm\mathbf{R}_i\mathbf{U}_i\} \subset G \times SU(2) \tag{5.23}$$

The two signs correspond to those of (5.17). Let K be the set of all such $\mathbf{U}_i$:

$$K = \{\pm\mathbf{U}_i | \mathbf{U} \to \mathbf{R}_i\} \tag{5.24}$$

K and G^* are isomorphic, $K \cong G^*$. The relations between the four groups are seen in figure (5.2).

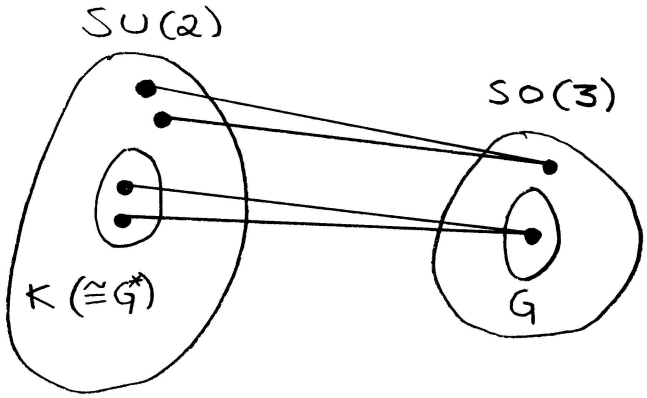

Figure 5.2: Group relations

5.2.3 Improper Rotations

The improper groups, G', are either direct products

$$G' = G \times C_i \tag{5.25}$$

where $C_i = \{E, I\}$ contains the inversion, I, or isomorphic to G. In the first case G' contains I.

Example 3:

$$T_\mathrm{h} = T \times C_i = \{E, 3C_2, 4C_3, 3C_3^{-1}\}\mathbf{U}\{I, 3\sigma, 4S_6^{-1}, 4S_6\} \tag{5.26}$$

5.2.4 The Group $O(3)$

All improper groups, G', are subgroups of the group $O(3)$, which is obtained from $SO(3)$ by adding the improper rotation with $\det \mathbf{A} = \pm 1$.

If the inversion is included, it commutes with all other elements, $\begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$ being a multiple of $I = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$.

5.2.5 The Double Group is a Symmetry Group of the Dirac Equation.

Consider

$$h_{\mathrm{D}} = c\underline{\alpha} \cdot \underline{p} + \beta mc^2 + V \tag{5.27}$$

with

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{pmatrix} \tag{5.28}$$

and define the transformation matrix for the i^{th} rotation as

$$\mathbf{U}_i^{\mathrm{D}} = \begin{pmatrix} \mathbf{U}_i & 0 \\ 0 & \mathbf{U}_i \end{pmatrix} \tag{5.29}$$

Consider the group

$$G_{\mathrm{D}}^* = \{\pm \mathbf{R}_i \mathbf{U}_i^{\mathrm{D}}\} \cong G^* \tag{5.30}$$

This rotation transforms the kinetic energy to

$$(\mathbf{R}_i \mathbf{U}_i^{\mathrm{D}})\underline{\alpha} \cdot \underline{p} (\mathbf{R}_i \mathbf{U}_i^{\mathrm{D}})^\dagger = \begin{pmatrix} 0 & \mathbf{U}\sigma_k \mathbf{U}^\dagger \\ \mathbf{U}\sigma_k \mathbf{U}^\dagger & 0 \end{pmatrix} \mathbf{R} p_k \mathbf{R}^\dagger \tag{5.31}$$

$$= \alpha_l p_m \mathbf{A}_{lk} \mathbf{A}_{mk} = \alpha_m p_m = \underline{\alpha} \cdot \underline{p}$$

Thus the kinetic energy remains invariant under the double group operations.

5.2.6 The Element $\bar{E}$

Bethe (1929) arrived at the double-group concept by including in the group the new element $\bar{E}$, corresponding to rotations through 2π, for which

$$\bar{E}^2 = E \tag{5.32}$$

Example 4:

$$C_2 = \{E, C_2\}, \quad C_2^* = \{E, C_2, \bar{E}, \bar{C}_2 = \bar{E}C_2\} \tag{5.33}$$

In general

$$\bar{A} = \bar{E}A \tag{5.34}$$

5.2.7 Elements of Double Groups

In the multiplication table of a double group the new, barred, operations are mixed with the others. For rotations around the same axis

$$C_2\mathbf{R}_i = \begin{cases} \bar{\mathbf{R}}_j, & \text{for } 0 < \mathbf{R}_i < \pi \\ \mathbf{R}_j, & \text{for } -\pi < \mathbf{R}_i < 0 \end{cases} \tag{5.35}$$

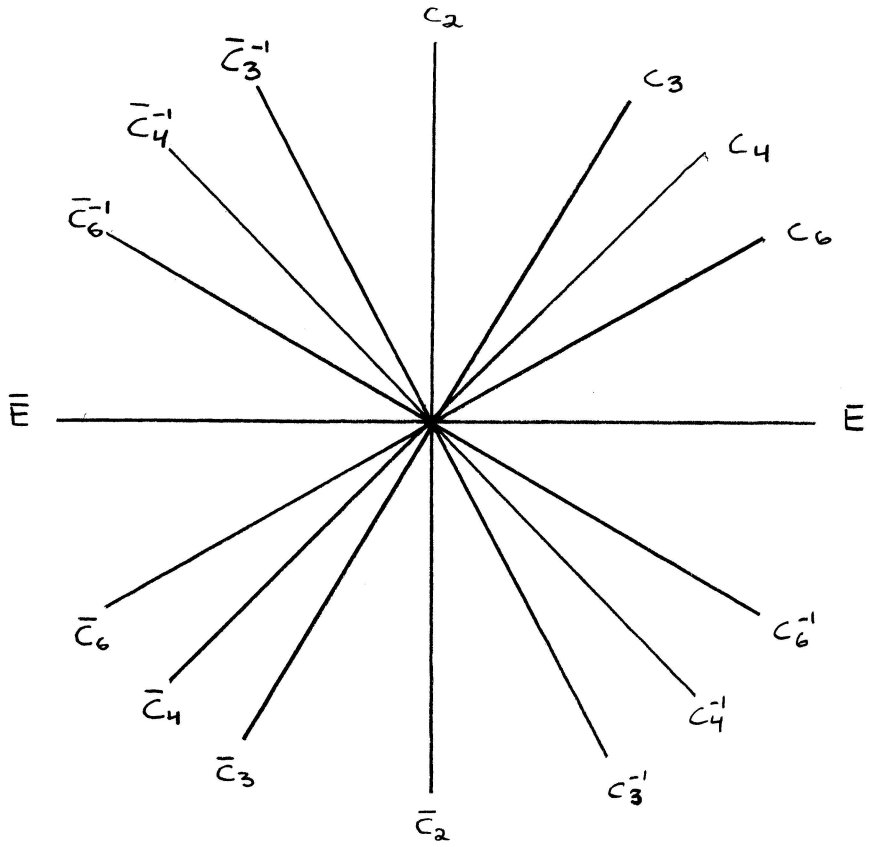

Figure 5.3: Multiplication of rotations in a double group. From **Koster** *et al.* (1963).

The properties of the inversion are

$$II = E, \quad I\bar{I} = \bar{E}, \quad \bar{I}\bar{I} = E, \quad I\mathbf{R}_i = \mathbf{S}_j, \quad I\bar{\mathbf{R}}_j = \bar{\mathbf{S}}_j \tag{5.36}$$

As

$$\sigma_{\rm h} = IC_2 \tag{5.37}$$

we get

$$\sigma_{\rm h} C_2 = IC_2^2 = I\bar{E} = \bar{I} \neq I\,! \tag{5.38}$$
$$\sigma\sigma = (IC_2)(IC_2) = C_2^2 = \bar{E} \tag{5.39}$$
$$\sigma\bar{\sigma} = \sigma\sigma\bar{E} = \bar{E}\bar{E} = E \Rightarrow \tag{5.40}$$
$$\sigma^{-1} = \bar{\sigma}\,! \tag{5.41}$$

Also,

$$\{E, \sigma\} \otimes \{E, \bar{E}\} = \{E, \sigma, \bar{E}, \underbrace{\sigma^{-1}}_{\bar{\sigma}}\} \tag{5.42}$$
$$\sigma\sigma^{-1} = E \tag{5.43}$$
$$\sigma\bar{E} = \bar{\sigma} \tag{5.44}$$
$$\bar{E}\sigma^{-1} = \bar{E}\bar{\sigma} = \bar{E}\bar{E}\sigma = \sigma \tag{5.45}$$

For a rotation C_n,

$$C_n^n = \bar{E}, \quad C_n^{2n} = E \Rightarrow \tag{5.46}$$
$$\bar{E}C_n = C_n^n C_n = C_n C_n^n = C_n\bar{E} \tag{5.47}$$

Furthermore,

$$C_2\mathbf{R}_i = \begin{cases} \bar{\mathbf{R}}_j, & \text{for } 0 < \mathbf{R}_i < \pi \\ \mathbf{R}_j, & \text{for } -\pi < \mathbf{R}_i < 0 \end{cases} \tag{5.35}$$

$\Rightarrow$

$$\sigma_{\rm h}\mathbf{R}_i = IC_2\mathbf{R}_i = \begin{cases} \bar{\mathbf{S}}_i, & \text{for } 0 < \mathbf{R}_i < \pi \\ \mathbf{S}_i, & \text{for } -\pi < \mathbf{R}_i < 0 \end{cases} \tag{5.48}$$

5.2.8 Irreducible Representations ("irreps")

The order of the double group, $\#G^* = 2(\#G) = 2g$. Hence its irreps must satisfy

$$\sum_{i=1}^{r} n_i^2 = 2g \tag{5.49}$$

r being the number of irreps for G^*.

Example 5:

Both C_{2h} and C_{2v} have $g = 4$. They satisfy (5.49) by

$$C_{2h} : (1^2 + 1^2 + 1^2 + 1^2) + (1^2 + 1^2 + 1^2 + 1^2) = 8 \tag{5.50}$$

$$C_{2v} : (1^2 + 1^2 + 1^2 + 1^2) + 2^2 = 8 \tag{5.51}$$

$$\Gamma_1 \quad \ldots \quad \Gamma_4 \quad \Gamma_5$$

Thus all relativistic molecular orbitals of C_{2v} belong to the same irrep, Γ_5. Note that, for an even number of electrons, the *total* electronic symmetry still belongs to $\Gamma_1 - -\Gamma_4$, J being an integer (for higher symmetries, when applicable).

The reps of G are also reps of G^*

We have a 2-to-1 mapping from G^* to G and a n-to-1 mapping from G to reps $\{\Gamma\}$. Thus we have a $2n$-to-1 mapping from G^* to a single matrix of the representation Γ.

The original irreps of G have basis functions, $|jm\rangle$, with integer j. For their characters

$$\chi^{(i)}(\bar{\mathbf{A}}) = \chi^{(i)}(\mathbf{A}) \tag{5.52}$$

The additional irreps of G^* have basis functions $|jm\rangle$ with *half-integer* j. For them

$$\chi^{(i)}(\bar{\mathbf{A}}) = -\chi^{(i)}(\mathbf{A}) \tag{5.53}$$

Here i is the irrep in question.

5.2.9 Classes

The number of classes equal the number of irreps,

$$k = r \tag{5.54}$$

The character of the rep, spanned by the functions $|jm\rangle$, $m = -j, \ldots, j$, becomes

$$\chi^{(j)}(\omega) = \frac{\sin(j + \frac{1}{2})\omega}{\sin \frac{1}{2}\omega} \tag{5.55}$$

For $\omega = \pi$ and half-integer j, $\chi^{(j)}(\pi) = 0$. Therefore $\chi^{(j)}(\pi + 2\pi) = -\chi^{(j)}(\pi) = 0$, as well. Thus $\chi(C_2) = \chi(\bar{C}_2) \Rightarrow C_2$ and $\bar{C}_2$ may or may not belong to the same class.

For $n \geq 2$, one always has $\chi(C_n) = -\chi(\bar{C}_n)$ and here C_n and $\bar{C}_n$ belong to different classes.

5.2.10 Theorem of Opechowski

Opechowski (1940): C_n and $\bar{C}_n$ belong to to different classes if $n \geq 2$. For $n = 2$, C_2 and $\bar{C}_2$ belong to the same class if and only if the group contains another C_2, perpendicular to the C_2 considered, or if it contains a symmetry plane, σ, containing the C_2 considered.

Similarly, S_2 and $\bar{S}_2$ belong to different classes except if there is a σ, containing the C_2 axis, or another C_2 perpendicular to it.

The reflections σ and $\bar{\sigma}$ belong to different classes, except if there exists another σ, perpendicular to it, or a C_2 in the plane σ.

Example 6:
Character table and basis functions for the group D^*_{3h}:

D_{3h}	E	$\bar{E}$	σ_h $\bar{\sigma}_h$	$2C_3$	$2\bar{C}_3$	$2S_3$	$2\bar{S}_3$	$3C_2'$ $3\bar{C}_2'$	$3\sigma_v$ $3\bar{\sigma}_v$	Time inv.	Bases
Γ_1	1	1	1	1	1	1	1	1	1	a	R
Γ_2	1	1	1	1	1	1	1	-1	-1	a	S_z
Γ_3	1	1	-1	1	1	-1	-1	1	-1	a	zS_z
Γ_4	1	1	-1	1	1	-1	-1	-1	1	a	z
Γ_5	2	2	-2	-1	-1	1	1	0	0	a	$(S_x\text{-i}S_y)$, $\text{-}(S_x\text{-i}S_y)$
Γ_6	2	2	2	-1	-1	-1	-1	0	0	a	$\Gamma_3 \times \Gamma_5$
Γ_7	2	-2	0	1	-1	$\sqrt{3}$	$-\sqrt{3}$	0	0	c	$\phi(\frac{1}{2}, \text{-}\frac{1}{2})$, $\phi(\frac{1}{2}, \frac{1}{2})$
Γ_8	2	-2	0	1	-1	$-\sqrt{3}$	$\sqrt{3}$	0	0	c	$\Gamma_7 \times \Gamma_3$
Γ_9	2	-2	0	-2	2	0	0	0	0	c	$\phi(\frac{3}{2}, \text{-}\frac{3}{2})$, $\phi(\frac{3}{2}, \frac{3}{2})$

We observe that: $D_{3h} = \{E; 2C_3; 3C_2; \sigma_h; 2S_3; 3\sigma_v\}$

a. The order of D_{3h} is 12. The representations Γ_7–Γ_9 satisfy (5.49) by $2^2 + 2^2 + 2^2 = 12$.

b. The elements σ_h and $\bar{\sigma}_h$ belong to the same class because the C_2 axes lie in the σ_h plane. The "vertical" σ_v and $\bar{\sigma}_v$ belong to the same class both because of the C_2 and because of σ_h. C_2 and $\bar{C}_2$ belong to the same class because of either σ_h or one σ_v.

c. By Opechowski's theorem, C_3 and $\bar{C}_3$ belong to different classes. So do S_3 and $\bar{S}_3$. The element $\bar{E}$ forms its own class. This gives the three classes, imposed by $k = r$, eq. (5.54).

5.3 Construction of Relativistic MO:s

5.3.1 Projection Operators

The functions ϕ_λ^k, belonging to row λ of irrep k can be projected out from an arbitrary function by using the *projection operator*

$$P_{\lambda\nu}^{(k)} = \frac{n_k}{2g} \sum_A \Gamma_{\lambda\nu}^{(k)}(A)^* \mathcal{O}_A \tag{5.56}$$

Here $g = \#G$, n_k= dimension if irrep k, $\underline{\underline{\Gamma}}^{(k)}(A)$ is the representation matrix for group element A in the irrep k and $\underline{\underline{\mathcal{O}}}_A$ is the symmetry operation A.

One can first operate on an arbitrary function, ψ, by $P_{\lambda\lambda}^{(k)}$ obtaining

$$\phi_\lambda^k = P_{\lambda\lambda}^{(k)} \psi \tag{5.57}$$

One then obtains the partners of ϕ_λ^k by

$$\phi_\nu^k = P_{\nu\lambda}^{(k)} \phi_\lambda^k \tag{5.58}$$

Suppose now that the basis functions in our LCAO-MO consist of atomic spinors,

$$\psi_{jlm} = \frac{1}{r} \begin{pmatrix} P(r)\chi_{jlm}(\vartheta,\phi) \\ \mathrm{i}Q(r)\chi_{j\bar{l}m}(\vartheta,\phi) \end{pmatrix} \tag{5.59}$$

The bispinors $\chi_{jlm} = |l\tfrac{1}{2}jm\rangle$ transform under the rotations according to $\mathrm{D}^j(\alpha\beta\gamma)$.

As noted by **Rosén** and **Ellis** (1979), then the result of the projection operator becomes

$$P_{\nu\lambda}^{(k)} \chi_{jlm} = \frac{n_k}{2g} \sum_{A=1}^{2g} \Gamma_{\nu\lambda}^{(k)}(A)^* \sum_{m'=-j}^{j} \chi_{jlm'} (-)^{l\tau_a} \mathrm{D}_{m'm}^j(\alpha\beta\gamma) \tag{5.60}$$

Here $\tau_A = 1$ if A contains the inversion operator I and $\tau_A = 0$ otherwise. Note that, outside the symmetry center, the result of $\mathcal{O}_A \psi_{jlm}^t$ may be shifted from the atom t to other, symmetry-equivalent atoms.

Example 7:

(Pyykkö and **Toivanen** 1977): Relativistic symmetry orbitals for the double group $D_{3\mathrm{h}}$. The origin is situated at the symmetry center. The variable s is +1 for $\kappa < 0$ and -1 for $\kappa > 0$.

irrep	even l	odd l
Γ_7	$\begin{cases} \lvert lj\frac{1}{2}\rangle \\ s\lvert lj\text{-}\frac{1}{2}\rangle \end{cases}$	$\begin{cases} \lvert lj\text{-}\frac{5}{2}\rangle \\ \text{-}s\lvert lj\frac{5}{2}\rangle \end{cases}$
Γ_8	$\begin{cases} \lvert lj\text{-}\frac{5}{2}\rangle \\ s\lvert lj\frac{5}{2}\rangle \end{cases}$	$\begin{cases} \lvert lj\frac{1}{2}\rangle \\ \text{-}s\lvert lj\text{-}\frac{1}{2}\rangle \end{cases}$
Γ_9	$\begin{cases} \lvert lj\frac{3}{2}\rangle \\ \text{-}s\lvert lj\text{-}\frac{3}{2}\rangle \end{cases}$	$\begin{cases} \lvert lj\text{-}\frac{3}{2}\rangle \\ s\lvert lj\frac{3}{2}\rangle \end{cases}$

Latest program: **J. Meyer** *et al.* (1996).

5.3.2 Coupling Constant Method

For the required coupling constants, see **Koster** *et al.* (1963) or **Altmann** and **Herzig** (1994). One now first constructs the non-relativistic MO:s, u^i_λ, and then couples them with the spin functions,

$$\begin{cases} \vartheta^j_{\text{-}1/2} = |\frac{1}{2}, -\frac{1}{2}\rangle = \beta \\ \vartheta^j_{1/2} = |\frac{1}{2}, \quad \frac{1}{2}\rangle = \alpha, \end{cases} \tag{5.61}$$

using the *coupling constants* or ***Clebsch-Gordan*** *coefficients* $u^{ij,k}_{\lambda\mu,\nu}$:

$$\phi^k_\nu = \sum_{\lambda\mu} u^{ij,k}_{\lambda\mu,nu} u^i_\lambda v^j_\nu. \tag{5.62}$$

The non-relativistic funcitons u^i_λ are expressed in the $|lm_l\rangle$ representation. From that $|lm_l\frac{1}{2}m_s\rangle$ representation, (5.62), one can pass to $|l\frac{1}{2}jm_j\rangle$ representation using the Clebsch-Gordan coefficients

	j	
m_s	$l+\frac{1}{2}$	$l-\frac{1}{2}$
$\frac{1}{2}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$	$-\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}$
$-\frac{1}{2}$	$\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$

This method will automatically give the correct phase between the $j = l \pm \frac{1}{2}$ components.

5.4 Time Reversal

5.4.1 Non-relativistic Case

Time reversal is obtained by complex conjugating the wave function,

$$T\psi = \psi^* \equiv K\psi. \tag{5.63}$$

Let $T = K$ (complex conj.), $[H, T] = 0$.

$$\big(\mathrm{i}\hbar\frac{\partial}{\partial t}\big)\psi = H\psi \Rightarrow \big(\mathrm{i}\hbar\frac{\partial}{\partial(-t)}\big)\psi^* = T(H\psi) = H(T\psi) = H\psi^* \tag{5.64}$$

5.4.2 Inclusion of Spin

Now

$$T\psi(\underline{r}, s) = (-\mathrm{i})^{2s}\psi(\underline{r}, -s)^*. \tag{5.65}$$

This operator T is *antilinear*,

$$T(a_1\psi_1 + a_2\psi_2) = a_1^*T\psi_1 + a_2^*T\psi_2, \tag{5.66}$$

and *antiunitary*,

$$\langle T\psi_1|T\psi_2\rangle = \langle\psi_2|\psi_1\rangle. \tag{5.67}$$

5.4.3 n-electron Wave Functions

We now can write T as the product of (antiunitary) operator K and a unitary operator U. Possible choices are

$$T = \underbrace{\sigma_{1y}\sigma_{2y}\ldots\sigma_{ny}}_{u}K, \qquad \Big|\sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \tag{5.68}$$

or

$$T\psi(\underline{r}_1, s_1, \ldots, \underline{r}_n, s_n) = (-\mathrm{i})^2(s_1 + \ldots + s_n)\psi(\underline{r}_1, -s_1, \ldots, \underline{r_n}, -s_n)^*. \tag{5.69}$$

5.4.4 Kramers' Theorem

For a system with an odd number of electrons, ψ and $T\psi$ are orthogonal.

In general, for arbitrary ψ and ϕ,

$$\langle T\psi|T\phi\rangle = \langle UK\psi|UK\phi\rangle = \langle K\psi|K\psi\rangle = \langle\phi|\psi\rangle \tag{5.70}$$

$$\Rightarrow \langle T\psi|\psi\rangle \overset{(5.70)}{=} \langle T\psi|T^2\psi\rangle \qquad \Big|T^2\psi = -\psi$$

$$= \langle T\psi| - \psi\rangle = -\langle T\psi|\psi\rangle = 0 \tag{5.71}$$

Note that the equivalence used, $T^2\psi = -\psi$, only holds for an odd number of electrons!

In the absence of magnetic field, all energy levels of such systems are at least doubly degenerate.

5.4.5 The Cases (a), (b) and (c) of Wigner (1932)

Consider the representation matrices, $\underline{\underline{\Gamma}}$, of a group G^*. The operator T either does or does not give new, degenerate functions, spanning the representation matrices $\underline{\underline{\Gamma}}^*$. The following possibilities exist:

a. $\underline{\underline{\Gamma}}$ may be real.

b. $\underline{\underline{\Gamma}}$ is complex and non-equivalent to $\underline{\underline{\Gamma}}^*$

c. $\underline{\underline{\Gamma}}$ is complex and equivalent to $\underline{\underline{\Gamma}}^*$

Then:

a. The degeneracy is doubled (from n for Γ to $2n$. The functions ψ and $T\psi$ belong to the same irrep.

b. The degeneracy is doubled. If ψ spans the irrep Γ, ψ^* spans another irrep Γ^*.

c. The degeneracy remains n.

According to the ***Frobenius-Schur*** *criterion*,

$$\sum_{A=1}^{2n} \chi^{(i)}(A^2) = \begin{cases} 2g, & \text{case (a)} \\ 0, & \text{case (b)} \\ -2g, & \text{case (b)} \end{cases} \tag{5.72}$$

5.4.6 Further Examples

Some other examples of time inversion are considered below.

Momentum

$$T(\underline{p}\psi) = T\Big((-\mathrm{i}\hbar\nabla)\psi\Big) = (\mathrm{i}\hbar\nabla)\psi^2 = -(\underline{p}T\psi \tag{5.73}$$

$$T\underline{p} = -\underline{p}T, \quad \text{the sign of } \underline{p} \text{ is inverted!} \tag{5.74}$$

$$T\underline{p}T^{-1} = -\underline{p} \tag{5.75}$$

Angular momentum

$$\underline{l} = \underline{r} \times \underline{p} \qquad \big|(5.74),\ \underline{p} \text{ changes sign under time inversion}$$
$$\Rightarrow T\underline{l} = -\underline{l}T \tag{5.76}$$

Spin-orbit Hamiltonian

$$\begin{aligned} h_{\mathrm{SO}} &= \frac{\hbar^2}{4mc^2}(\underline{\sigma} \times \nabla V) \cdot \underline{p} \quad \big| \text{both } \underline{\sigma} \text{ and } \underline{p} \text{ change sign,} \\ &\to h_{\mathrm{SO}} \end{aligned} \tag{5.77}$$

Spin functions

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$T\alpha = (-\mathrm{i})^{2\cdot\frac{1}{2}}\beta \overset{(5.69)}{=} -\mathrm{i}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5.78}$$

$$T\beta = (-\mathrm{i})^{2\cdot(-\frac{1}{2})}\alpha = \ \mathrm{i}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{5.79}$$

$$\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \overset{(5.68)}{=} \ \mathrm{i}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \mathrm{i} \end{pmatrix}$$

$$\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ = \ -\mathrm{i}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\mathrm{i} \\ 0 \end{pmatrix}$$

Rösch: $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Rightarrow$

$$\begin{cases} T\alpha = \beta \\ T\beta = -\alpha \end{cases} \tag{5.80}$$

Eschrig:

$$T\Psi(\alpha) = \begin{pmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & 0 & 1 \\ & & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ cp_z/A \\ c(p_x + \mathrm{i}p_y)/A \end{pmatrix} \mathrm{e}^{-\mathrm{i}\underline{p}\cdot\underline{r}} \tag{5.81}$$

$$= \begin{pmatrix} 0 \\ -1 \\ -c(p_x - \mathrm{i}p_y)/A \\ cp_z/A \end{pmatrix} \mathrm{e}^{-\mathrm{i}\underline{p}\cdot\underline{r}} = -\mathrm{e}^{-\mathrm{i}\underline{p}\cdot\underline{r}} \begin{pmatrix} 0 \\ 1 \\ c(p_x - \mathrm{i}p_y)/A \\ -cp_z/A \end{pmatrix} = -\Psi(\beta)$$

Compare the above with (3.45). Still, $T^2 = -\underline{\underline{1}}$

External electromagnetic fields
Invert the currents,

$$\underline{j} \to -\underline{j}, \quad \underline{A} \to -\underline{A} \tag{5.82}$$

5.5 Quaternions

For complex numbers $z = x + \mathrm{i}y = (x, y)$ both addition:

$$(x, y) + (X, Y) = (x + X, y + Y) \tag{5.83}$$

and multiplication:

$$(x, y)(X, Y) = (xX - yY, xY + yX) \tag{5.84}$$

are defined. The multiplication is

a. *commutative*, $z_1 z_2 = z_2 z_1$

b. *associative*, $a(bc) = (ab)c$

c. *norm-conserving*, $|z_1 z_2| = |z_1||z_2|$, $|z|^2 = x^2 + y^2$

Question: Can (a)-(c) be satisfied for more than two ordered components?

Answer: Hamilton and **Cayley** (1843): For *quaternions* $q = (a, b, c, d)$, with $|q|^2 = a^2 + b^2 + c^2 + d^2$, a non-commutative product with (c) can be found.

6 Molecular Orbital Methods

6.1 Semiempirical methods

As a general philosophy, the purpose of semiempirical calculations is to produce good insight, not to produce bad numbers.

6.1.1 Extended Hückel Methods

Include all valence electrons. Diagonalize an effective, one-electron Hamiltonian

$$[\mathbf{H} - E_i\mathbf{S}]\mathbf{c}_i = 0 \tag{6.1}$$

$$h_{ab} = (KS_{ab}/2)[h_{aa} + h_{bb}] \tag{6.2}$$
$$h_{aa} = \alpha_a \tag{6.3}$$

K is the **Wolfsberg-Helmholz** constant and has a value of 1.75.

The overlap matrix **S** us evaluated using (single- or multiple-ζ) Slater orbitals:

$$S_{ab} = \langle a|b\rangle \tag{6.4}$$

The total energy is approximated by the eigenvalue sum

$$E_T = \sum_{i=1}^{\text{occ}} E_i \tag{6.5}$$

Carry out **Mulliken** population analysis. Can be made charge-iterative.

Historical origins:

- Mulliken 1949, Wolfsberg and Helmholz 1952; Eq. (6.2)
- **Longuet-Higgins** *et al.* 1954, **Lipscomb** *et al.* 1961
- **Hoffmann 1963**

Use in "quasirelativistic mode":

- Hoffmann school, including band structures.

For a review on relativistic semiempirical methods, see Pyykkö (1988b).

6.1.2 Zero Differential Overlap Approximation

Include all valence electrons. Start from the Hartree-Fock equations

$$\begin{aligned} \mathbf{F}\psi_1 &= [T + V_n + V_c + V_x]\psi_i \\ &= E_i\psi_i + \sum_{j\neq i} E_{ij}\psi_j \end{aligned} \tag{6.6}$$

and set systematically all products of two AOs,

$$\phi_A\phi_b = \delta_{ab}\phi_a^2 \tag{6.7}$$

in the potentials V_c and V_x, and in $\mathbf{S}$.

The total energy is calculated as

$$E_{\mathrm{T}} = \langle\Psi|F|\Psi\rangle \tag{6.8}$$

with the single-determinant wave function

$$\Psi = |\psi_1(1)\overline{\psi_1}(2)\ldots| \tag{6.9}$$

Various versions

- CNDO (Complete Neglect of Differential Overlap): Include only the integrals

$$\begin{aligned} &\langle a_A(1)b_B(1)|r_{12}^{-1}|c_C(2)d_D(2)\rangle \\ &\quad = \delta_{AB}\delta_{ab}\delta_{CD}\delta_{cd}(aa|cc) \end{aligned} \tag{6.10}$$

- INDO (Intermediate NDO): Add all one-center integrals. Exchange introduced.

Quasirelativistic ZDO models

- **Boča** 1987–, **Rösch/Zerner** *et al.* 1987–. Use relativistic data.
- **M. J. S. Dewar** *et al.* (1985): Empirical MNDO or AM1 parameters for Hg, Pb. See review by J. J. P. Stewart, *J. Computer-Aided Mol. Design* **4** (1990) 1-105.
- See Table 7.7 of RTAM III (2000).
- **Thiel** and **Voityuk** (1996): MNDO for d-orbitals, Cl–I, Zn–Hg.

6.1.3 Inclusion of Spin-Orbit Splitting

Many calculations since late 1960'ies, see refs. 69-102 in P. Pyykkö, *Methods in Comp. Chem.* **2** (1988) 137-226. Simplest alternative: include the one-center SO term,

$$h_{\mathrm{SO}} = \sum_{g}^{\text{centers}} \zeta_g \mathbf{1} \cdot \mathbf{s} \tag{6.11}$$

This obviously doubles the Hamiltonian matrix (for α and β spin).

ZDO Methods with SO Splitting

- Boča (1989, 1990ab): Relativistic CNDO/1. Applications on di- and triatomics, $Au_6^{2+}, \ldots$
- **Böhm** *et al.* (1989): Relativistic ZDO. Fe_n, $n = 12 - 14$.
- **Kotzian** *et al.* (1989, 1991, 1992), Kotzian and Rösch (1991, 1992): Relativistic INDO/S – CI ('S' for spectroscopic). Diatomic hydrides and oxides. LnO. $[Ce(H_2O)_6]^{3+}$.
- **Jahns** *et al.* (1992): INDO/1 ('1' for ground state) $[(Cp_2^*M)_2C_6H_4]$, M = Sc, Lu. *Agostic* M...H attraction reproduced.
- **Minaev** *et al.* (1989): Relativistic MINDO/3 – CI. Diatomic hydride systems. Intensity borrowing.
- **Roszak** and **Lipiński** (1992): INDO. Triplet – singlet lifetimes of pyridine etc.

6.1.4 Relativistic Extended Hückel (REX)

- Introduce $|lsjm\rangle$ basis with separate energy parameters, α_a and radial parameters ζ_a for the SO-split atomic states.
- Derive these parameters from *ab initio* atomic calculations, thus extrapolating them to the molecular domain (Lohr and Pyykkö 1979). R/NR!
- Rösch (1983) treated the *Kramers degeneracy* algebraically (QATREX)
- **Larsson** and Pyykkö (1986): Charge iterations added (ITEREX)
- Original default parameters (in 'DEFPAR') acceptable for main-group elements. Realistic parameters for the actinides in Pyykkö, Laakkonen and Tatsumi, *Inorg. Chem.* (28) (1989) 1801.
- Always carry out a *sensitivity study* with respect to parameters or try to confirm the ideas found by better methods.

Some Results by REX

- Implement the relativistic theory of nuclear spin-spin coupling constants. New terms. See Pyykkö and **Wiesenfield** (1981).
- Symmetry rules for the spin-spin coupling tensor, **J**, **A. D. Buckingham** *et al.* (1982).
- Transparent interpretation of the *heavy-atom chemical shift* in ^{1}HX NMR, Pyykkö, **Görling** and Rösch (1987). "HAHA" (1987).
- The *6p-hole theory* of actinyl NQCC, Larsson and Pyykkö (1986).
- Colors of $BiPh_5$ and $PbCl_6^{2-}$ attributed to the relativistic stabilization of the 6s a_1 LUMO (**Schmuck** *et al.* (1990), **El-Issa** *et al.* (1991)).
- The SO splitting of the OsO_4 valence $3t_2$ MO attributed to 5p semicore character, in most models. Pyykkö, **Li**, **Bastug**, **Fricke** and **Kolb** (1993).

6.2 One-Electron Molecules

6.2.1 The Hamiltonian

$$h = T + V, \quad V = -\frac{Z_1}{r_1} - \frac{Z_2}{r_2}. \tag{6.12}$$

6.2.2 Possible Coordinate Systems

a) *Elliptical*:

$$\begin{cases} \xi = (r_1 + r_2)/R, & 1 \le \xi < \infty \\ \eta = (r_1 - r_2)/R, & -1 \le \eta \le 1 \end{cases} \tag{6.13}$$

$$\begin{cases} r_1 = \dfrac{R}{2}(\xi + \eta) \\ r_2 = \dfrac{R}{2}(\xi - \eta) \end{cases} \tag{6.14}$$

Constant ξ: ellipses, constant η: hyperbolas.

Third variable: ϕ.

$$\mathrm{d}V = \frac{R^3}{8}(\xi^2 - \eta^2)\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\phi \tag{6.14'}$$

Then

$$h\psi = E\psi \tag{6.15}$$

$$\psi = \mathrm{e}^{\mathrm{i}m\phi} f(\xi, \eta) \tag{6.16}$$

$$V(\xi, \eta) = -\frac{2Z_1}{R(\xi + \eta)} - \frac{2Z_2}{R(\xi - \eta)} \tag{6.17}$$

The Schrödinger equation:

$$\Big[-\frac{2}{R^2(\xi^2 - \eta^2)}\Big(\frac{\partial}{\partial\xi}(\xi^2 - 1)\frac{\partial}{\partial\xi} + \frac{\partial}{\partial\eta}(1 - \eta^2)\frac{\partial}{\partial\eta} - m^2\Big(\frac{1}{\xi^2 - 1} + \frac{1}{1 - \eta^2}\Big)\Big) + V(\xi, \eta) - E\Big] f(\xi, \eta) = 0 \tag{6.18}$$

Ø. Burrau, *Mat. fys. Medd.* **7** (1927)

D. R. Bates, K. Ledsham, A. L. Stewart, *Phil. Trans. RSL* **246** (1953) 215.

See L. Laaksonen *et al.* (1983a).

Eq. (6.18) separates,

$$f(\xi, \eta) = L(\xi)M(\eta) \tag{6.19}$$

The eigenvalues E are obtained from the truncation of the series for $L(\xi)$.

b) *The variable q*:

$$\eta = 1 - N \ln q, \quad 0 \le q \le 1 \tag{6.20}$$

c) *The variable t*:

$$t = \ln \xi \tag{6.21}$$

d) *Prolate spheroidal*:

The 2D-program uses the variables (μ, ν):

$$\begin{cases} \xi = \cosh\mu, & 0 \le \mu < \infty \\ \eta = \cos\nu, & 0 \le \nu \le \pi \end{cases} \tag{6.22}$$

$$\begin{cases} r_1 = \frac{1}{2}R(\cosh\mu + \cos\nu) \\ r_2 = \frac{1}{2}R(\cosh\mu - \cos\nu) \end{cases} \tag{6.23}$$

$$
\begin{aligned}
\mathrm{d}V &= \frac{R^3}{8}\sinh\mu\sin\nu(\cosh^2\mu-\cos^2\nu)\mathrm{d}\mu\mathrm{d}\nu\mathrm{d}\phi && (6.24)\\
\psi &= \mathrm{e}^{\mathrm{i}m\phi}f(\mu,\nu) && (6.25)\\
\frac{\nabla^2\psi}{\mathrm{e}^{\mathrm{i}m\phi}} &= \frac{4}{R^2(\cosh^2\mu-\cos^2\nu)}\Big[\frac{\partial^2}{\partial\mu^2}+\coth\mu\frac{\partial}{\partial\mu}+\frac{\partial^2}{\partial\nu^2}\\
&\quad+\cot\nu\frac{\partial}{\partial\nu}-m^2\big(\frac{1}{\sinh^2\mu}+\frac{1}{\sin\nu}\big)\Big]f(\mu,\nu) && (6.26)
\end{aligned}
$$

e)*Kolb's coordinates* (s,t):

See C. Düsterhöft, L. Yang, D. Heinemann, D. Kolb, *Chem. Phys. Lett.* **229** (1994) 667.

$$
\begin{aligned}
\eta &= 4\sinh^4 s+1, \qquad && 0\le s<\infty\\
\xi &= \cos t[1+\frac{1}{2}\sin^2 t], && 0\le t\le\pi
\end{aligned}
\tag{6.27}
$$

[Already in L. Yang *et al.*, *Chem. Phys. Lett.* **178** (1991) 213]

6.2.3 Transformation of the Dirac Equation

D. Sundholm, P. Pyykkö, L. Laaksonen, *Phys. Scripta* **36** (1987) 400.
D. Sundholm, *Chem. Phys. Lett.* **223** (1994) 469.

Consider the Dirac hamiltonian

$$
h_{\mathrm{D}} = \begin{pmatrix} V & c\underline{\sigma}\cdot\underline{p} \\ c\underline{\sigma}\cdot\underline{p} & V-2c^2 \end{pmatrix} \tag{6.28}
$$

$$
\underline{\sigma}\cdot\underline{p} = \begin{pmatrix} p_z & p_x-\mathrm{i}p_y \\ p_x+\mathrm{i}p_y & -p_z \end{pmatrix} \tag{6.29}
$$

Let

$$
\Psi = \begin{pmatrix}\psi_{\mathrm{L}}\\ \psi_{\mathrm{S}}\end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{L}}(\mu,\nu)\\ \mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{L}}(\mu,\nu)\\ \mathrm{i}\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{S}}(\mu,\nu)\\ \mathrm{i}\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{S}}(\mu,\nu)\end{pmatrix} \tag{6.30}
$$

The lower equation from (6.28) gives

$$
\begin{aligned}
c\underline{\sigma}\cdot\underline{p}\psi_{\mathrm{L}}+(V-2c^2)\psi_{\mathrm{S}} &= E\psi_{\mathrm{S}} \Rightarrow\\
\psi_{\mathrm{S}} &= \frac{c\underline{\sigma}\cdot\underline{p}}{2c^2+E-V}\psi_{\mathrm{L}} && (6.31)
\end{aligned}
$$

This can be substituted in the upper equation,

$$\begin{aligned}(V-E)\psi_{\mathrm{L}} &= -c\underline{\sigma}\cdot\underline{p}\psi_{\mathrm{S}} \\ &= -c\underline{\sigma}\cdot\underline{p}(2c^2+E-V)^{-1}c\underline{\sigma}\cdot\underline{p}\psi_{\mathrm{L}} \end{aligned} \tag{6.32}$$

Letting

$$f(\underline{r}) = \frac{c^2}{2c^2+E-V}, \tag{6.33}$$

$$[\underline{\sigma}\cdot\underline{p}f(\underline{r})\underline{\sigma}\cdot\underline{p}+V-E]\psi_{\mathrm{L}} = 0 \tag{6.34}$$

$$\underline{p} = -\mathrm{i}\nabla, \quad p^2 = -\nabla^2, \quad \mathrm{i}p = \nabla$$

Recalling

$$(\underline{\sigma}\cdot\underline{A})(\underline{\sigma}\cdot\underline{B}) = \underline{A}\cdot\underline{B} + \mathrm{i}\underline{\sigma}\cdot(\underline{A}\times\underline{B}) \qquad [3.88]$$

$$\begin{aligned}\underline{\sigma}\cdot\underline{p}f(\underline{r})\underline{\sigma}\cdot\underline{p} &= \underline{\sigma}\cdot\big[(\underbrace{\underline{p}f}_{\underline{A}})\underline{\sigma}\cdot\underbrace{\underline{p}}_{\underline{B}} + f\underline{p}(\underline{\sigma}\cdot\underline{p})\big] \\ &= (\underline{p}f)\cdot\underline{p} + \mathrm{i}\underline{\sigma}\cdot(\underline{p}f\times\underline{p}) + f\big[\ \underbrace{(\underline{\sigma}\cdot\underline{p})(\underline{\sigma}\cdot\underline{p})}_{p^2+\mathrm{i}\underline{\sigma}\cdot\underbrace{(p\times p)}_{=0}}\ \big] \\ &= fp^2 + (\underline{p}f)\cdot\underline{p} + \mathrm{i}\underline{\sigma}\cdot(\underline{p}f\times p) \\ &= -f\nabla^2 - \nabla f\cdot\nabla - \mathrm{i}\underline{\sigma}\cdot(\nabla f\times\nabla) \end{aligned} \tag{6.35}$$

In an orthogonal coordinate system,

$$\nabla f = \hat{e}_1\frac{1}{h_1}\frac{\partial f}{\partial q_1} + \hat{e}_2\frac{1}{h_2}\frac{\partial f}{\partial q_2} + \hat{e}_3\frac{1}{h_3}\frac{\partial f}{\partial q_3} \tag{6.36}$$

(The distance, $\mathrm{d}s_i = h_i\mathrm{d}q_i$). Here

$$\begin{aligned}h_\mu = h_\nu &= \frac{R}{2}\sqrt{(\cosh^2\mu - \cos^2\mu)} \\ h_\phi &= \frac{R}{2}\sinh\mu\sin\nu \end{aligned} \tag{6.37}$$

where all the hs are weight functions.

Then

$$\begin{aligned}-\nabla f\cdot\nabla\psi &= -\frac{1}{h_\mu^2}\frac{\partial f}{\partial\mu}\frac{\partial\psi}{\partial\mu} - \frac{1}{h_\nu^2}\frac{\partial f}{\partial\nu}\frac{\partial\psi}{\partial\nu} - \frac{1}{h_\phi^2}\frac{\partial f}{\partial\phi}\frac{\partial\psi}{\partial\phi} \\ &= -\frac{4}{R^2(\cosh^2\mu-\cos^2\nu)}\big[\frac{\partial f}{\partial\mu}\frac{\partial\psi}{\partial\mu} + \frac{\partial f}{\partial\nu}\frac{\partial\psi}{\partial\nu}\big] \\ &\quad -\frac{4}{R^2\sinh^2\mu\sin^2\nu}\,\frac{\partial f}{\partial\phi}\frac{\partial\psi}{\partial\phi} \end{aligned} \tag{6.38}$$

$$-f\nabla^2 = -\frac{4f}{R^2(\cosh^2\mu - \cos^2\nu)}\Big[\frac{\partial^2}{\partial\mu^2} + \frac{\cosh\mu}{\sinh\mu}\frac{\partial}{\partial\mu} + \frac{\partial^2}{\partial\nu^2} + \frac{\cos\nu}{\sin\nu}\frac{\partial}{\partial\nu} + \frac{\cosh^2\mu - \cos^2\nu}{\sinh^2\mu\sin^2\nu}\frac{\partial^2}{\partial\phi^2}\Big] \tag{6.39}$$

What, then, is $-\mathrm{i}\underline{\sigma}\cdot(\underbrace{\nabla f\times\nabla\psi}_{\underline{D}})$?

On one hand,

$$\nabla f\times\nabla\psi = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial y} & \frac{\partial\psi}{\partial z} \end{vmatrix} \tag{6.40}$$

(all hs equal 1 in cartesian coordinates)

On the other hand, denoting $f_x \equiv \frac{\partial f}{\partial x}$,

$$\begin{aligned} \nabla f\times\nabla\psi &= \begin{vmatrix} \hat{e}_\mu & \hat{e}_\nu & \hat{e}_\phi \\ \frac{1}{h_\mu}f_\mu & \frac{1}{h_\nu}f_\nu & \frac{1}{h_\phi}f_\phi \\ \frac{1}{h_\mu}\psi_\mu & \frac{1}{h_\nu}\psi_\nu & \frac{1}{h_\phi}\psi_\phi \end{vmatrix} \\ &= \hat{e}_\mu\frac{1}{h_\nu h_\phi}(f_\nu\psi_\phi - f_\phi\psi_\nu) + \hat{e}_\nu\frac{1}{h_\phi h_\mu}(f_\phi\psi_\mu - f_\mu\psi_\phi) \\ &\quad + \hat{e}_\phi\frac{1}{h_\mu h_\nu}(f_\mu\psi_\nu - f_\nu\psi_\mu) \end{aligned} \tag{6.41}$$

We will need the cartesian components of this determinant,

$$\underline{D} = D_1\hat{e}_1 + D_2\hat{e}_2 + D_3\hat{e}_3 \tag{6.42}$$

for the scalar product with

$$\underline{\sigma} = \sigma_1\hat{e}_1 + \sigma_2\hat{e}_2 + \sigma_3\hat{e}_3 \tag{6.43}$$

$$\begin{cases} \sigma_\mu = \frac{1}{h_\mu}(\sigma_x\frac{\partial x}{\partial\mu} + \sigma_y\frac{\partial y}{\partial\mu} + \sigma_z\frac{\partial z}{\partial\mu}) \\ \sigma_\nu = \frac{1}{h_\nu}(\sigma_x\frac{\partial x}{\partial\nu} + \sigma_y\frac{\partial y}{\partial\nu} + \sigma_z\frac{\partial z}{\partial\nu}) \\ \sigma_\phi = \frac{1}{h_\phi}(\sigma_x\frac{\partial x}{\partial\phi} + \sigma_y\frac{\partial y}{\partial\phi} + \sigma_z\frac{\partial z}{\partial\phi}) \end{cases} \tag{6.44}$$

$$
\begin{aligned}
&-\mathrm{i}\underline{\sigma}\cdot(\nabla f\times\nabla\psi)= \\
-\frac{\mathrm{i}}{h_\mu h_\nu h_\phi}\ &\Big[\ (\sigma_x\frac{\partial x}{\partial\mu}+\sigma_y\frac{\partial y}{\partial\mu}+\sigma_z\frac{\partial z}{\partial\mu})(f_\nu\psi_\phi-f_\phi\psi_\nu) \\
&+\ (\sigma_x\frac{\partial x}{\partial\nu}+\sigma_y\frac{\partial y}{\partial\nu}+\sigma_z\frac{\partial z}{\partial\nu})(f_\phi\psi_\mu-f_\mu\psi_\phi) \\
&+\ (\sigma_x\frac{\partial x}{\partial\phi}+\sigma_y\frac{\partial y}{\partial\phi}+\sigma_z\frac{\partial z}{\partial\phi})(f_\mu\psi_\nu-f_\nu\psi_\mu)\ \Big]
\end{aligned}
\tag{6.45}
$$

From

$$
\begin{cases}
x=\dfrac{R}{2}\sinh\mu\sin\nu\cos\phi \\
y=\dfrac{R}{2}\sinh\mu\sin\nu\sin\phi \\
z=\dfrac{R}{2}\cosh\mu\cos\nu\quad (z\text{ is in the molecular axis direction})
\end{cases}
\tag{6.46}
$$

$$
\begin{cases}
\dfrac{\partial x}{\partial\mu}=\dfrac{R}{2}\cosh\mu\sin\nu\cos\phi, & \dfrac{\partial x}{\partial\nu}=\dfrac{R}{2}\sinh\mu\cos\nu\cos\phi, \\
& \dfrac{\partial x}{\partial\phi}=\dfrac{R}{2}\sinh\mu\sin\nu\sin\phi \\
\dfrac{\partial y}{\partial\mu}=\dfrac{R}{2}\cosh\mu\sin\nu\sin\phi, & \dfrac{\partial y}{\partial\nu}=\dfrac{R}{2}\sinh\mu\cos\nu\sin\phi, \\
& \dfrac{\partial y}{\partial\phi}=\dfrac{R}{2}\sinh\mu\sin\nu\cos\phi \\
\dfrac{\partial z}{\partial\mu}=\dfrac{R}{2}\sinh\mu\cos\nu, & \dfrac{\partial z}{\partial\nu}=-\dfrac{R}{2}\cosh\mu\sin\nu, \\
& \dfrac{\partial z}{\partial\phi}=0
\end{cases}
\tag{6.47}
$$

$$
\begin{aligned}
-\mathrm{i}\underline{\sigma}\cdot\underline{D}=-\frac{\mathrm{i}R/2}{h_\mu h_\nu h_\phi}\big[& \\
&(\sigma_x\cosh\mu\sin\nu\cos\phi+\sigma_y\cosh\mu\sin\nu\sin\phi+\sigma_z\sinh\mu\cos\nu)(f_\nu\psi_\phi-f_\phi\psi_\nu) \\
+&(\sigma_x\sinh\mu\cos\nu\cos\phi+\sigma_y\sinh\mu\cos\nu\sin\phi-\sigma_z\cosh\mu\sin\nu)(f_\phi\psi_\mu-f_\mu\psi_\phi) \\
+&(-\sigma_x\sinh\mu\sin\nu\sin\phi+\sigma_y\sinh\mu\sin\nu\cos\phi+0)(f_\mu\psi_\nu-f_\nu\psi_\mu)\big]
\end{aligned}
\tag{6.48}
$$

$$
\begin{cases}
-\sigma_x D_x = -\dfrac{\frac{R}{2}\sigma_x}{h_\mu h_\nu h_\phi}\Big[\cosh\mu\sin\nu\cos\phi(f_\nu\psi_\phi - f_\phi\psi_\nu) \\
\qquad\qquad + \sinh\mu\cos\nu\cos\phi(f_\phi\psi_\mu - f_\mu\psi_\phi) - \sinh\mu\sin\nu\sin\phi(f_\mu\psi_\nu - f_\nu\psi_\mu)\Big] \\
-\sigma_y D_y = -\dfrac{\frac{R}{2}\sigma_y}{h_\mu h_\nu h_\phi}\Big[\cosh\mu\sin\nu\sin\phi(f_\nu\psi_\phi - f_\phi\psi_\nu) \\
\qquad\qquad + \sinh\mu\cos\nu\sin\phi(f_\phi\psi_\mu - f_\mu\psi_\phi) - \sinh\mu\sin\nu\cos\phi(f_\mu\psi_\nu - f_\nu\psi_\mu)\Big] \\
-\sigma_z D_z = -\dfrac{\frac{R}{2}\sigma_z}{h_\mu h_\nu h_\phi}\Big[\sinh\mu\cos\nu(f_\nu\psi_\phi - f_\phi\psi_\nu) \\
\qquad\qquad - \cosh\mu\sin\nu(f_\phi\psi_\mu - f_\mu\psi_\phi)\Big]
\end{cases}
\tag{6.49}
$$

Recall that

$$
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
$$

$$
-\mathrm{i}\underline{\sigma}\cdot\underline{D} = -\frac{\mathrm{i}R/2}{h_\mu h_\nu h_\phi} \times
$$

$$
\begin{pmatrix}
\cosh\mu\sin\nu\cos\phi(f_\nu\psi_\phi - f_\phi\psi_\nu) & \\
+\sinh\mu\cos\nu\cos\phi(f_\phi\psi_\mu - f_\mu\psi_\phi) & \\
-\sinh\mu\sin\nu\sin\phi(f_\mu\psi_\nu - f_\nu\psi_\mu) & \\
\sinh\mu\cos\nu(f_\nu\psi_\phi - f_\phi\psi_\nu) & -\mathrm{i}\Big[\cosh\mu\sin\nu\sin\phi(f_\nu\psi_\phi - f_\phi\psi_\nu) \\
-\cosh\mu\sin\nu(f_\phi\psi_\mu - f_\mu\psi_\phi) & +\sinh\mu\cos\nu\sin\phi(f_\phi\psi_\mu - f_\mu\psi_\phi) \\
& +\sinh\mu\sin\nu\cos\phi(f_\mu\psi_\nu - f_\nu\psi_\mu)\Big] \\
\cosh\mu\sin\nu\cos\phi(f_\nu\psi_\phi - f_\phi\psi_\nu) & \\
+\sinh\mu\cos\nu\cos\phi(f_\phi\psi_\mu - f_\mu\psi_\phi) & \\
-\sinh\mu\sin\nu\sin\phi(f_\mu\psi_\nu - f_\nu\psi_\mu) & -\sinh\mu\cos\nu(f_\nu\psi_\phi - f_\phi\psi_\nu) \\
+\mathrm{i}\Big[\cosh\mu\sin\nu\sin\phi(f_\nu\psi_\phi - f_\phi\psi_\nu) & +\cosh\mu\sin\nu(f_\phi\psi_\mu - f_\mu\psi_\phi) \\
+\sinh\mu\cos\nu\sin\phi(f_\phi\psi_\mu - f_\mu\psi_\phi) & \\
+\sinh\mu\sin\nu\cos\phi(f_\mu\psi_\nu - f_\nu\psi_\mu)\Big] &
\end{pmatrix}
$$

It can be noted that the matrix in the above equation is of the form

$$
\begin{pmatrix} a & u - \mathrm{i}[v] \\ u + \mathrm{i}[v] & -a \end{pmatrix}.
$$

$$-\mathrm{i}\underline{\sigma}\cdot\underline{D} = -\frac{\mathrm{i}R/2}{\frac{R}{2}^3(\cosh^2\mu-\cos^2\nu)\sinh\mu\sin\nu}\times \tag{6.50}$$
$$\begin{pmatrix} \sinh\mu\cos\nu(f_\nu\psi_\phi - f_\phi\psi_\nu) & \cosh\mu\sin\nu(\cos\psi-\mathrm{i}\sin\phi)(f_\nu\psi_\phi-f_\phi\psi_\nu) \\ -\cosh\mu\sin\nu(f_\phi\psi_\mu - f_\mu\psi_\phi) & +\sinh\mu\cos\nu(\cos\phi-\mathrm{i}\sin\phi)(f_\phi\psi_\mu-f_\mu\psi_\phi) \\ & -\sinh\mu\sin\nu(\sin\phi+\mathrm{i}\cos\phi)(f_\mu\psi_\nu-f_\nu\psi_\mu) \\ \cosh\mu\sin\nu(\cos\psi+\mathrm{i}\sin\phi)(f_\nu\psi_\phi-f_\phi\psi_\nu) & -\sinh\mu\cos\nu(f_\nu\psi_\phi-f_\phi\psi_\nu) \\ +\sinh\mu\cos\nu(\cos\phi+\mathrm{i}\sin\phi)(f_\phi\psi_\mu-f_\mu\psi_\phi) & +\cosh\mu\sin\nu(f_\phi\psi_\mu-f_\mu\psi_\phi) \\ -\sinh\mu\sin\nu(\sin\phi-\mathrm{i}\cos\phi)(f_\mu\psi_\nu-f_\nu\psi_\mu) & \end{pmatrix}$$

Now note that

$$f_\mu = \frac{c^2(-1)}{(2c^2+E-V)^2}(-\frac{\partial V}{\partial\mu}),\quad f_\nu = \frac{c^2(-1)}{(2c^2+E-V)^2}(-\frac{\partial V}{\partial\nu}), \tag{6.51}$$
$$f_\phi \propto \frac{\partial V}{\partial\phi} = 0,$$

$$\cos\phi\mp\mathrm{i}\sin\phi = \begin{cases}\mathrm{e}^{-\mathrm{i}\phi}\\ \mathrm{e}^{\mathrm{i}\phi}\end{cases},\quad \sin\phi\pm\mathrm{i}\cos\phi = \begin{cases}\mathrm{i}\mathrm{e}^{-\mathrm{i}\phi}\\ -\mathrm{i}\mathrm{e}^{\mathrm{i}\phi}\end{cases} \tag{6.52}$$

$$-I\underline{\sigma}\cdot\underline{D} = -\frac{4\mathrm{i}}{R^2(\cosh^2\mu-\cos^2\nu)\sinh\mu\sin\nu}\times\frac{c^2}{(2c^2+E-V)^2}\times \tag{6.53}$$
$$\begin{pmatrix} \sinh\mu\cos\nu\dfrac{\partial V}{\partial\nu}\psi_\phi & \cosh\mu\sin\nu\,\mathrm{e}^{-\mathrm{i}\phi}\dfrac{\partial V}{\partial\nu}\psi_\phi \\ +\cosh\mu\sin\nu\dfrac{\partial V}{\partial\mu}\psi_\phi & -\sinh\mu\cos\nu\,\mathrm{e}^{-\mathrm{i}\phi}\dfrac{\partial V}{\partial\mu}\psi_\phi \\ & -\sinh\mu\sin\nu(\mathrm{i}\mathrm{e}^{-\mathrm{i}\phi})\Big(\dfrac{\partial V}{\partial\mu}\dfrac{\partial}{\partial\nu}-\dfrac{\partial V}{\partial\nu}\dfrac{\partial}{\partial\mu}\Big)\psi \\ \cosh\mu\sin\nu\,\mathrm{e}^{\mathrm{i}\phi}\dfrac{\partial V}{\partial\nu}\psi_\phi & -\sinh\mu\cos\nu\dfrac{\partial V}{\partial\nu}\psi_\phi \\ -\sinh\mu\cos\nu\,\mathrm{e}^{\mathrm{i}\phi}\dfrac{\partial V}{\partial\mu}\psi_\phi & -\cosh\mu\sin\nu\dfrac{\partial V}{\partial\mu}\psi_\phi \\ +\sinh\mu\sin\nu(\mathrm{i}\mathrm{e}^{\mathrm{i}\phi})\Big(\dfrac{\partial V}{\partial\mu}\dfrac{\partial}{\partial\nu}-\dfrac{\partial V}{\partial\nu}\dfrac{\partial}{\partial\mu}\Big)\psi & \end{pmatrix}$$

Recalling the Ansatz (6.30),

$$\psi = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{L}}(\mu,\nu) \\ \mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{L}}(\mu,\nu) \\ \mathrm{i}\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{S}}(\mu,\nu) \\ \mathrm{i}\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{S}}(\mu,\nu) \end{pmatrix},$$
$$\psi_\phi = \begin{pmatrix} \mathrm{i}(m-\frac{1}{2})\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{L}}(\mu,\nu) \\ \mathrm{i}(m+\frac{1}{2})\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{L}}(\mu,\nu) \\ -(m-\frac{1}{2})\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{S}}(\mu,\nu) \\ -(m+\frac{1}{2})\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{S}}(\mu,\nu) \end{pmatrix} = \frac{\partial\psi}{\partial\phi} \tag{6.54}$$

the operation $-\mathrm{i}\underline{\sigma}\cdot\underline{D}$ gives for the upper row:

$$\begin{aligned} A\mathrm{i}\Big\{ &\Big[\sinh\mu\cos\nu\frac{\partial V}{\partial\nu} + \cosh\mu\sin\nu\frac{\partial V}{\partial\mu}\Big]\mathrm{i}(m-\frac{1}{2})\psi_1^{\mathrm{L}}\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi} \\ &+\Big[\cosh\mu\sin\nu\frac{\partial V}{\partial\nu} - \sinh\mu\cos\nu\frac{\partial V}{\partial\mu}\Big]\mathrm{e}^{-\mathrm{i}\phi}\mathrm{i}(m+\frac{1}{2})\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{L}} \\ &-\sinh\mu\sin\nu\,\mathrm{i}\mathrm{e}^{-\mathrm{i}\phi}\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\nu} - \frac{\partial V}{\partial\nu}\frac{\partial}{\partial\mu}\Big)\psi_2^{\mathrm{L}}\Big\} \end{aligned}$$

$$A = -\frac{4c^2}{R^2(\cosh^2\mu - \cos^2\nu)\sinh\mu\sin\nu(2c^2+E-V)^2} \tag{6.56}$$

and for the lower row:

$$\begin{aligned} A\mathrm{i}\Big\{ &\Big[\cosh\mu\sin\nu\frac{\partial V}{\partial\nu} - \sinh\mu\cos\nu\frac{\partial V}{\partial\mu}\Big]\mathrm{e}^{\mathrm{i}\phi}\mathrm{i}(m-\frac{1}{2})\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\psi_1^{\mathrm{L}} \\ &+\sinh\mu\sin\nu\,\mathrm{i}\mathrm{e}^{\mathrm{i}\phi}\mathrm{e}^{\mathrm{i}(m-\frac{1}{2})\phi}\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\nu} - \frac{\partial V}{\partial\nu}\frac{\partial}{\partial\mu}\Big)\psi_1^{\mathrm{L}} \\ &-\Big[\sinh\mu\cos\nu\frac{\partial V}{\partial\nu} + \cosh\mu\sin\nu\frac{\partial V}{\partial\mu}\Big]\mathrm{i}(m+\frac{1}{2})\mathrm{e}^{\mathrm{i}(m+\frac{1}{2})\phi}\psi_2^{\mathrm{L}}\Big\} \end{aligned} \tag{6.57}$$

Dividing away the $\mathrm{e}^{\mathrm{i}\phi}$-factors

$$\Rightarrow -\mathrm{i}\underline{\sigma}\cdot\underline{D} = \tag{6.58}$$
$$\begin{pmatrix} -A\Big\{\Big[\sinh\mu\cos\nu\dfrac{\partial V}{\partial\nu}+\cosh\mu\sin\nu\dfrac{\partial V}{\partial\mu}\Big](m-\frac{1}{2})\psi_1^{\mathrm{L}} \\ \quad+\Big[\Big(\cosh\mu\sin\nu\dfrac{\partial V}{\partial\nu}-\sinh\mu\cos\nu\dfrac{\partial V}{\partial\mu}\Big)(m+\frac{1}{2}) \\ \quad\quad-\sinh\mu\sin\nu\Big(\dfrac{\partial V}{\partial\mu}\dfrac{\partial}{\partial\nu}-\dfrac{\partial V}{\partial\nu}\dfrac{\partial}{\partial\mu}\Big)\Big]\psi_2^{\mathrm{L}}\Big\} \\ \\ -A\Big\{\Big[\Big(\cosh\mu\sin\nu\dfrac{\partial V}{\partial\nu}-\sinh\mu\cos\nu\dfrac{\partial V}{\partial\mu}\Big)(m-\frac{1}{2}) \\ \quad+\sinh\mu\sin\nu\Big(\dfrac{\partial V}{\partial\mu}\dfrac{\partial}{\partial\nu}-\dfrac{\partial V}{\partial\nu}\dfrac{\partial}{\partial\mu}\Big)\Big]\psi_1^{\mathrm{L}} \\ \quad-\Big[\sinh\mu\cos\nu\dfrac{\partial V}{\partial\nu}+\cosh\mu\sin\nu\dfrac{\partial V}{\partial\mu}\Big](m+\frac{1}{2})\psi_2^{\mathrm{L}}\Big\} \end{pmatrix}$$

Using the definition of A, (6.56),

$$\Rightarrow -\mathrm{i}\underline{\sigma}\cdot\underline{D} = +\frac{4}{R^2(\cosh^2\mu-\cos^2\nu)}\,\frac{f}{2c^2+E-V}\times \tag{6.59}$$
$$\begin{pmatrix} \Big[\dfrac{\cos\nu}{\sin\nu}\dfrac{\partial V}{\partial\nu}+\dfrac{\cosh\mu}{\sinh\mu}\dfrac{\partial V}{\partial\mu}\Big](m-\frac{1}{2})\psi_1^{\mathrm{L}} \\ +\Big[\Big(\dfrac{\cosh\mu}{\sinh\mu}\dfrac{\partial V}{\partial\nu}-\dfrac{\cos\nu}{\sin\nu}\dfrac{\partial V}{\partial\mu}\Big)(m+\frac{1}{2}) \\ -\Big(\dfrac{\partial V}{\partial\mu}\dfrac{\partial}{\partial\nu}-\dfrac{\partial V}{\partial\nu}\dfrac{\partial}{\partial\mu}\Big)\Big]\psi_2^{\mathrm{L}} \\ \\ \Big[\Big(\dfrac{\cosh\mu}{\sinh\mu}\dfrac{\partial V}{\partial\nu}-\dfrac{\cos\nu}{\sin\nu}\dfrac{\partial V}{\partial\nu}\Big)(m-\frac{1}{2}) \\ +\Big(\dfrac{\partial V}{\partial\mu}\dfrac{\partial}{\partial\nu}-\dfrac{\partial V}{\partial\nu}\dfrac{\partial}{\partial\mu}\Big)\Big]\psi_1^{\mathrm{L}} \\ -\Big[\dfrac{\cos\nu}{\sin\nu}\dfrac{\partial V}{\partial\nu}+\dfrac{\cosh\mu}{\sinh\mu}\dfrac{\partial V}{\partial\mu}\Big](m+\frac{1}{2})\psi_2^{\mathrm{L}} \end{pmatrix}$$

Combining this result with the *diagonal kinetic energy terms* (6.38)–(6.39) and $(V-E)\psi$, noting that in (6.38)

$$-\nabla f\cdot\nabla\psi = -\frac{4}{R^2(\cosh^2\mu-\cos^2\nu)}\,\frac{f}{2c^2+E-V}\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\mu}-\frac{\partial V}{\partial\nu}\frac{\partial}{\partial\nu}\Big)\psi$$

the upper equation finally becomes

$$
\begin{aligned}
-\frac{4f}{R^2(\cosh^2\mu-\cos^2\nu)}\Bigg\{\Bigg[&+\frac{\partial^2}{\partial\mu^2}+\frac{\cosh\mu}{\sinh\mu}\,\frac{\partial}{\partial\mu}+\frac{\partial^2}{\partial\nu^2}+\frac{\cos\nu}{\sin\nu}\,\frac{\partial}{\partial\nu}\\
&-\frac{\cosh^2\mu-\cos^2\nu}{\sinh^2\mu\sin^2\nu}(m-\frac{1}{2})^2\\
&+\frac{1}{2c^2+E-V}\Big[\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\mu}-\frac{\partial V}{\partial\nu}\frac{\partial}{\partial\nu}\Big)\\
&-\Big(\frac{\cos\nu}{\sin\nu}\,\frac{\partial V}{\partial\nu}+\frac{\cosh\mu}{\sinh\mu}\,\frac{\partial V}{\partial\mu}\Big)(m-\frac{1}{2})\Big]\Bigg]\psi_1^{\mathrm{L}}\\
&-\frac{1}{2c^2+E-V}\Big[\Big(\frac{\cosh\mu}{\sinh\mu}\,\frac{\partial V}{\partial\nu}-\frac{\cos\nu}{\sin\nu}\,\frac{\partial V}{\partial\mu}\Big)(m+\frac{1}{2})\\
&-\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\nu}-\frac{\partial V}{\partial\nu}\frac{\partial}{\partial\mu}\Big)\Big]\psi_2^{\mathrm{L}}\Bigg\}\\
&+(V-E)\psi_1^{\mathrm{L}}=0
\end{aligned}
\tag{6.60}
$$

and the lower equation becomes

$$
\begin{aligned}
&-\frac{4f}{R^2(\cosh^2\mu-\cos^2\nu)}\times\\
&\Bigg\{-\frac{1}{2c^2+E-V}\Big[\Big(\frac{\cosh\mu}{\sinh\mu}\,\frac{\partial V}{\partial\nu}-\frac{\cos\nu}{\sin\nu}\,\frac{\partial V}{\partial\mu}\Big)(m-\frac{1}{2})\\
&\qquad+\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\nu}-\frac{\partial V}{\partial\nu}\frac{\partial}{\partial\mu}\Big)\Big]\psi_1^{\mathrm{L}}\\
&+\frac{1}{2c^2+E-V}\Big[\Big(+\frac{\cos\nu}{\sin\nu}\,\frac{\partial V}{\partial\nu}+\frac{\cosh\mu}{\sinh\mu}\,\frac{\partial V}{\partial\mu}\Big)(m+\frac{1}{2})\\
&\qquad+\Big(\frac{\partial V}{\partial\mu}\frac{\partial}{\partial\mu}+\frac{\partial V}{\partial\nu}\frac{\partial}{\partial\nu}\Big)\Big]\psi_2^{\mathrm{L}}\\
&\quad+\Big[\frac{\partial^2}{\partial\mu^2}+\frac{\cosh\mu}{\sinh\mu}\,\frac{\partial}{\partial\mu}+\frac{\partial^2}{\partial\nu^2}+\frac{\cos\nu}{\sin\nu}\,\frac{\partial}{\partial\nu}\\
&\qquad-\frac{\cosh^2\mu-\cos^2\nu}{\sinh^2\mu\sin^2\nu}(m+\frac{1}{2})^2\Big]\psi_2^{\mathrm{L}}\Bigg\}\\
&+(V-E)\psi_2^{\mathrm{L}}=0
\end{aligned}
\tag{6.61}
$$

7 Pseudopotentials

7.1 Introduction

Heavy elements are the nightmare of theoreticians.

- Many electrons
- Many orbitals
- Relativistic dynamics

Solution: Find a *pseudopotential* with a ground state which is the lowest needed valence orbital.

- **Local PP** (the same for all values of l)
- **Semi-local PP** (different radial part for different values of l):,

$$V_l = \mathcal{P}_l V_l(r) \mathcal{P}_l, \tag{7.1}$$

$$\mathcal{P}_l = \sum_{m=-l}^{l} |lm\rangle\langle lm|. \tag{7.2}$$

Works.

7.2 A bit of history

- **H. Hellmann**, "Quantenchemie" (1937), p. 111– . The *Zusatzpotential* for alkali metals

$$U(r) = -\frac{1}{r} + \frac{A}{r} \mathrm{e}^{-2\kappa r} \tag{7.3}$$

- **P. Gombás**, *Z. Phys.* **94** (1935) (?) 473,479.
"Pseudopotentiale", Springer, Wien (1967).
"Ein Besetzungsverbotpotential".

- A historical wandering: **Phillips-Kleinman**, Phys. Rev. **116** (1959) 287.

$$(T + V)\psi_i = \varepsilon_i \psi_i, \tag{7.4}$$

$$V(r) = \frac{(\varepsilon_i - T)\psi_i(r)}{\psi_i(r)}. \tag{7.5}$$

Right ε_i, wrong norm of ψ.

- Fix: **shape-consistent PP**, **norm-conserving PP**. A discussion about which term was in use first is reproduced in figure (7.1) (P. Pyykkö, private communication to **R. Nieminen**, May 13, 1994).

- An own pseudo-potential for the spin-orbit interaction:

$$V^{\mathrm{SO}}(r) = \sum_{l=1}^{l_{\max}} \frac{2\Delta V_l(r)}{2l+1} \mathcal{P}_l \underline{l} \cdot \underline{s} \mathcal{P}_l. \tag{7.6}$$

$$\Delta V_l(r) = V_{l,l+\frac{1}{2}}(r) - V_{l-\frac{1}{2}}(r) \tag{7.7}$$

$$= \sum_k \Delta A_{lk} \mathrm{e}^{-a_{lk} r^2} \tag{7.8}$$

P.A. Christiansen, **Y.S. Lee**, **K.S. Pizer** *J. Chem. Phys.* **71** (1979) 4445.
P. Hafner, **W.H.E. Schwarz** *Chem. Phys. Lett.* **65** (1979) 537.

- The electric polarizability of the ionic core, and the core-valence correlation energy.

$$V_{\mathrm{CPP}} = -\frac{1}{2} \sum_\lambda \alpha_\lambda (\bar{f}_\lambda)^2 \tag{7.9}$$

Here α_λ is the polarizability of core λ and $\bar{f}_\lambda$ the electric field seen by the core λ.
P. Fuentealba, Thesis, Stuttgart (1984)
W. Müller, **J. Flesch**, **W. Meyer**, *J. Chem. Phys.* **80** (1984) 3297, 3311.

7.3 Where to get pseudopotentials

- **USA**:
  - K.S. Pitzer
  - Los Alamos (Hay & Wadt)
  - P.A. Christiansen
  - W. Ermler
  - W.J. Stevens (ECP=PP)
- **Canada, Sweden, Spain**: Huzinaga, Wahlgren, nodal model potentials.
- **Germany**: Stuttgart (Dolg, Stoll, Schwerdtfeger, ...) "Energy consistent PP".

  `www.theochem.uni-stuttgart.de`
- **France**: Toulouse (Daudley, Teichteil, ...)

F: pyykko@cc.helsinki.fi
T: rnieminen@csc.fi
D: 13 May, 1994
S: History of pseudopotentials

Dear Risto,

As you may recall we had some time ago a friendly
"who was first" discussion on the history of the
"shape-consistent" or "norm-conserving" pseudopotentials.
The former name is used in chemistry and the latter in physics;
the common feature of both is that the pseudo wave function
reproduce the full wave function in the valence region. (i)

At the workshop of Pseudopotentials of the REHE
programme of ESF in Toulouse on
Monday, I asked the members of the conference how they thought
that these ideas were originally developed.

Apparently the first article was /1/ in 1975. The idea (i) is
clearly stated in Fig. 2 and pp. 288-.

According to Walter Ermler, who was at Berkeley in the late
1970'ies, there was communication between the chemists, like him
and the physicists, like Dr. A. Zunger, working with M. Cohen.
Anecdotally, "we were rather shouting at each other than
collaborating", though.

The first paper from Berkeley/Chemistry seems to be /2/. It quotes /1/.
The Acknowledgement of /2/ thanks A. Zunger for a preprint of /3/,
"in the revision to final form of this manuscript", and
also L. Kahn for private communications on his work.
Thus the first papers /2, 3/
and the Physics and Chemistry efforts indeed seem entirely parallel.
None of the papers /1-3/ uses any particular name for the idea.

In his review /4/ Kahn quotes for the 'shape-consistent'
idea (i), his eq. (33), also /5, 6/. Kahn et al. /7/ also clearly
discuss the idea (i) in their II.B.ii, without quoting /1/.

Hamann et al. /6/ have the title 'Norm-conserving pseudopotentials'
and quote /3, 5/ but not /1, 7/.

Conclusion: The work in the references /1, 7 5/ seems to preceed
that in ref. /3, 6/. The name 'norm-conserving' form 1979 may be
older than the name 'shape-consistent', whose exact origin I have not
found. I believe that it was used in the Los Alamos workshop on the
subject in 1981.

1. Ph. Durand and J.-C. Barthelat, Theor. Chim. Acta 38 (1975) 283-202.
2. P.A. Christiansen, Y.S. Lee and K.S. Pitzer, J. Chem. Phys. 71 (1979)
   4445-4450.
3. A. Zunger, J. Vac. Sci. Technol. 16 (1979) 1337-1438.
4. L. Kahn, Int. J. Quantum Chem. 25 (1984) 149-183.
5. A. Redondo, W.A. Goddard III and T.C. McGill, Phys. Rev. B 15 (1977)
   5038-5048. Check ref. 6-7 there.
6. D. R. Hamann, M. Schlueter and C. Chiang, Phys. Rev. Letters 43 (1979)
   1494-1497.
7. L.R. Kahn, P. Baybutt and D.G. Truhlar, J. Chem. Phys. 65 (1976) 3826-
   3853.

Figure 7.1: e-mail, May 13, 1994

8 On QED

8.1 Introduction

Underpinning the Dirac-level theory we have *quantum electrodynamics*, QED. It is the theory on virtual photon processes. The two simplest terms are denoted by the **Feynman** diagrams

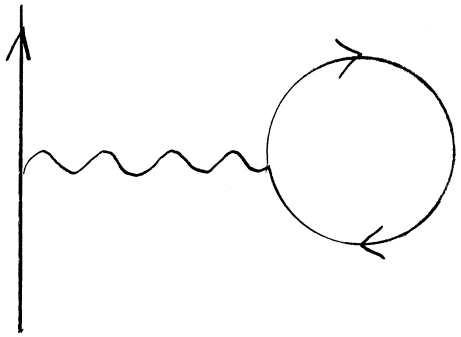

Figure 8.1: Vacuum polarization

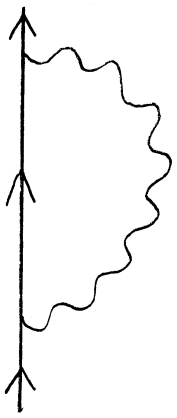

Figure 8.2: Self-energy

The lowest-order *vacuum polarization* (VP) term can be expressed in analytical form (**Uehling** 1935). It is independent of the nuclear charge, a property of the vacuum. Furthermore, it is attractive. The *self energy* (SE) is larger than the VP and of opposite sign. Both can be related (to lowest order) to $|\Psi(0)|^2$. The QED is a *covariant* theory.

8.2 Some formulas for vacuum polarization

$$V_n^{\text{eff}}(r) = -\frac{Z}{r}(1 + S(r)) = V_n + V_{\text{Ue}}, \tag{8.1}$$

$$S(r) = \frac{2\alpha}{3\pi}\int_1^\infty \exp\Big(-\frac{2r\chi}{\alpha}\Big)\Big(1+\frac{1}{\chi^2}\Big)\frac{\sqrt{\chi^2-1}}{\chi^2}d\chi \tag{8.2}$$

The Uehling function $S(r)$ can be fitted to the two-parameter expression

$$\begin{aligned} S(r) = \alpha\Big[&\exp(-d_1 r^2)C_1\left[\ln(\alpha/r) - C_2\right] \\ &+(1-\exp(-d_1 r^2))\Big(\frac{1}{C_3}\Big)\frac{\exp(-2r/\alpha)}{d_2(\frac{r}{\alpha})^{0.5}+(\frac{r}{\alpha})^{1.5}}\Big] \end{aligned} \tag{8.3}$$

Here the free parameters $d_1 = 0.678 \cdot 10^7$ and $d_2 = 1.4302$. The three constants are

$$C_1 = \frac{2}{2\pi}, \quad C_2 = \frac{5}{6} + C, \quad C_3 = 4\sqrt{\pi} \tag{8.4}$$

Also see figure (8.3).

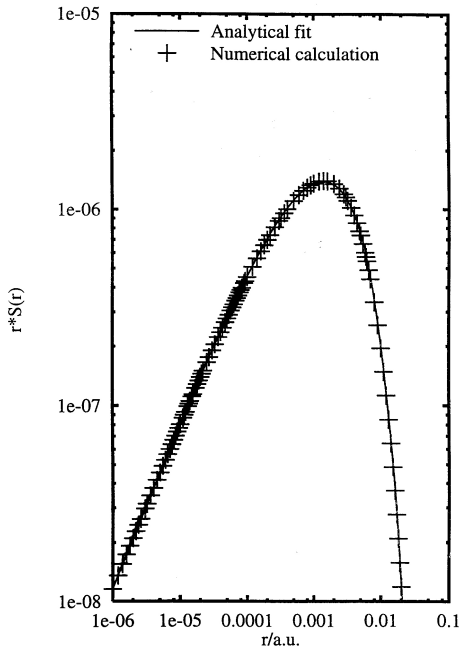

Figure 8.3: The Uehling function $rS(r)$

8.3 Some formulas for self-energy (vacuum fluctuation)

For light elements one can use the *density formula*

$$E_{\rm SE} = \frac{4Z\alpha^3}{3}\left[\ln\frac{1}{(\alpha Z)^2} - \ln\frac{2K_{n0}}{(\alpha Z)^2} + \frac{5}{6}\right]|\Psi(0)|^2 \tag{8.5}$$

Alternatively the total *Lamb shift* is obtained to lowest order from the density as

$$E_{\rm L} = \frac{4Z\alpha^3}{3}\left[\ln\frac{1}{(\alpha Z)^2} - \ln\frac{2K_{n0}}{(\alpha Z)^2} + \frac{19}{30}\right]|\Psi(0)|^2 \tag{8.6}$$

For heavy elements, evaluate $E_{\rm SE}$ from the $2s$ SE/VP ratio of **Johnson** and **Soff** (1985), The total valence-electron Lamb shift becomes

$$E_{\rm L} = \langle V_{\rm Ue}\rangle \frac{(E_{\rm SE} + E_{\rm VP})}{E_{\rm VP}} \tag{8.7}$$

For the valence electrons the order-of-magnitude of the Lamb shift remained unknown, beyond lithium, until the two papers

[1] P. Pyykkö, M. Tokman, L.N. Labzowsky, *Phys. Rev. A* **57** (1998) R 689

[2] L. Labzowsky, I. Goidenko, M. Tokman, P. Pyykkö, *Phys. Rev. A* **59** (1999) 2707

[1] used the density-based or the "ratio" methods. It also introduced a semiempirical "A-model" ('A' for the nuclear A-value). In [2], using spline-functions in a local model potential, the Feynman diagrams were actually calculated for the ns valence electrons of alkali and coinage metals. The results were very similar: For $Z \gtrapprox 50$, the Lamb shift is about -1% of the Dirac-level effects. Figures (8.4) and (8.5) give an idea of the size of the effects.

For lighter elements it is, relatively speaking, larger. A simple way to estimate it is to scale the Darwin term, as suggested in

[3] P. Pyykkö, K.G. Dyall, A.G. Császár, G. Tarczay, O.L. Polyansky and J. Tennyson, *Phys. Rev. A* **63** (2001), 024502, 1-4

$$E^{\rm Lamb}/E^{\rm Darwin} = \frac{8\alpha}{3\pi}\big[-2\ln(\alpha Z) - \ln X + \frac{19}{30}\big] \tag{8.8}$$

$$E^{\rm Lamb}/E^{\rm Darwin} = \frac{2\alpha F(Z\alpha)}{\pi} - \frac{8\alpha}{15\pi} \tag{8.9}$$

$$E^{\rm Lamb}/E^{\rm Darwin} = \frac{8\alpha}{3\pi}\ln(\frac{1}{Z\alpha}) \tag{8.10}$$

Equation (8.10) was introduced by **Bjorken** and **Drell** (1964). The results of using the above three equations can be studied in figure (8.6).

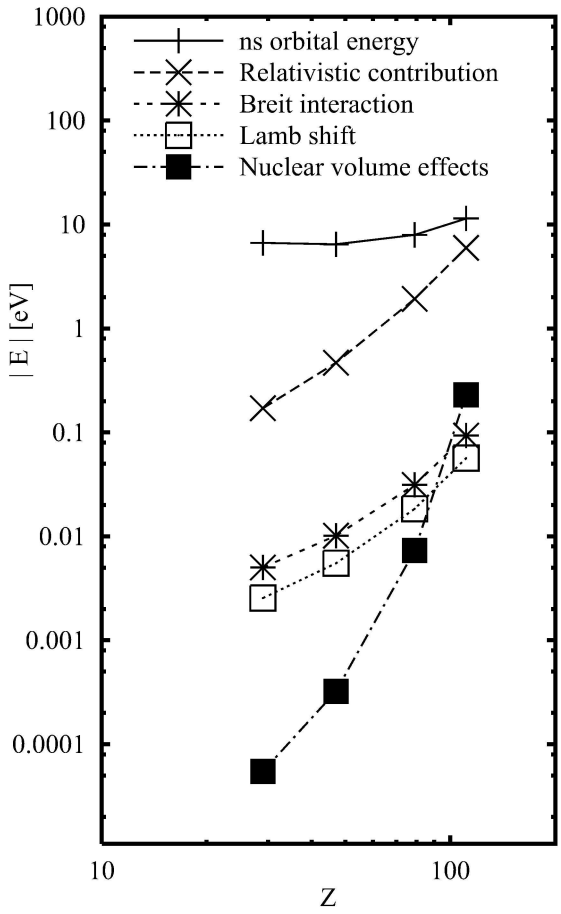

Figure 8.4: Dirac-Fock contributions for coinage metals. From P. Pyykkö, M. Tokman and L.N. Labzowsky, *Phys. Rev. A* **57** (1998) R689

Recall the Pauli Hamiltonian

$$h_p = -\frac{\alpha^2}{8}\underset{\sim}{p}^4 - \frac{\alpha^2}{8}\nabla^2 V - \frac{\alpha^2}{4}\underset{\sim}{\sigma}\cdot(\nabla V \times \underset{\sim}{p}) \tag{8.11}$$

where for a Coulomb potential

$$\nabla^2 V = -4Z\pi\delta(\underset{\sim}{r}) \tag{8.12}$$

X is related to the **Bethe logarithm** K_{n0}:

$$X = 2K_{n0}/(\alpha Z)^2 \;\approx\; 11.77, 16.64, 15.93, \ldots \quad \text{for 1s, 2s, 3s, } \ldots \tag{8.13}$$

The $F(\alpha Z)$ is used to express the exact self energy as

$$E_1^{\mathrm{SE}} = \alpha^3 Z F(\alpha Z)\langle\delta(\underset{\sim}{r})\rangle \tag{8.14}$$

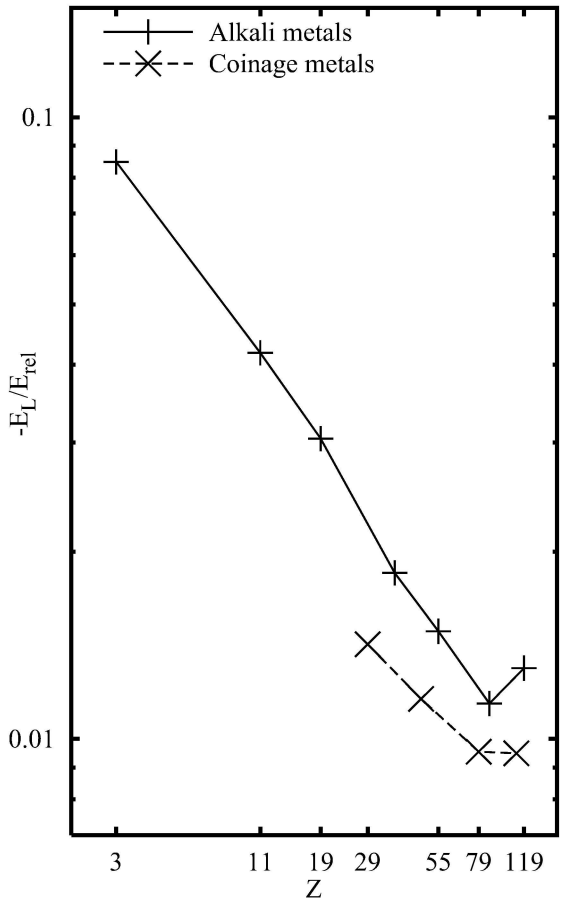

Figure 8.5: Ratios of the Lamb shift to the relativistic contribution. From L. Labzowsky, I. Goidenko, M. Tokman, P. Pyykkö, *Phys. Rev. A* **59** (1999) 2707

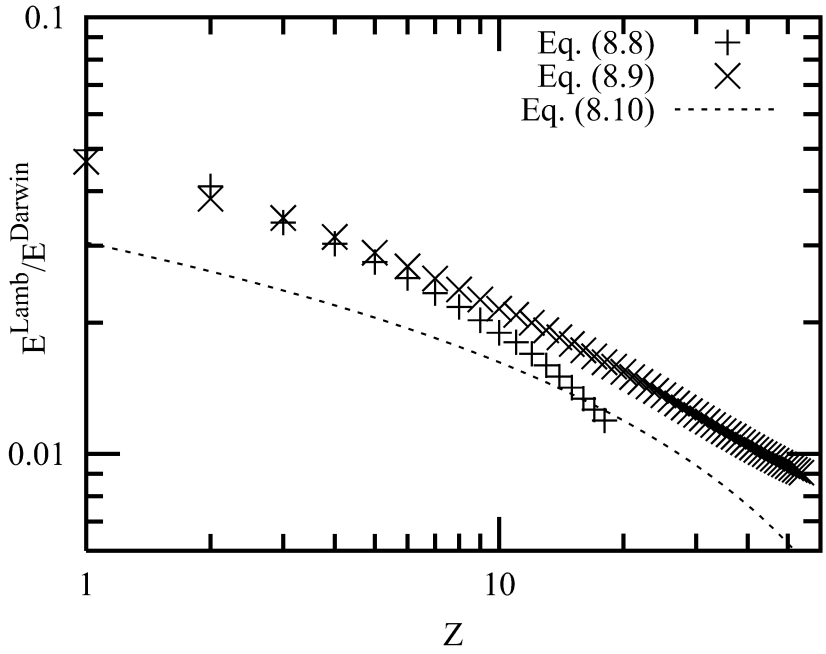

Figure 8.6: The ratio $E^{\mathrm{Lamb}}/E^{\mathrm{Darwin}}$. From P. Pyykkö *et al.* , *Phys. Rev. A* **63** (2001), 024502, 1-4

	$1S_{1/2}$
Binding energy E_B (point nucleus)	**-93459.89**
Corrections:	
Finite nuclear size	49.13
Self energy (order α)	196.68
VP: Uehling contribution	*-41.99*
VP: Wichmann-Kroll contribution	*1.79*
Total vacuum polarization (order α)	-40.20
SESE (2^{nd} order SE) (a) (b) (c)	
VPVP (2^{nd} order VP) (a) (ladder diagrams)	-0.07
VPVP (b) (Källén-Sabry contribution + h.o.)	-0.05
VPVP (c) (Källén-Sabry contribution)	-0.29
SEVP (a) (b) (c)	0.42
S(VP)E	0.05
Radiative recoil (estimate)	*0.00*
Reduced mass	*0.26*
Relativistic recoil	*0.08*
Total recoil	0.34
Nuclear polarization (*bottleneck for accuracy!*)	-0.02
Sum of corrections	205.99
Resulting total binding energy	**-93253.90**
Lamb shift (theory)	**205.73**
(experimental)	**202(8)**

Figure 8.7: One-electron Lamb shift contributions for $^{197}Au^{78+}$ (hydrogen-like gold). Energy values are given in eV. From T. Beier *et al.*, *Phys. Lett. A* **236** (1997) 329-338

9 On Transformed Hamiltonians

9.1 General

If the four-component Dirac equation is approximated by some two-component form, certain savings can be obtained. Typically, the small components are eliminated and two separate equations are introduced for atomic $j = l \pm \frac{1}{2}$ states. For molecules, one typically divides the Hamiltonian into a quasirelativistic (spin-orbit averaged) part and the spin-orbit part, diagonalized at some stage. Special attention must be given to variational stability.

One can follow these developments in Table 2.2 of RTAM I-III and the applications in Table 7.3 of RTAM III. For reviews, see **Hess** and **Marian** (2000a).

9.2 The Foldy-Wouthuysen transformation

Foldy and **Wouthuysen** (1950) presented a systematic procedure for decoupling the large and small components of the Dirac equation,

$$\begin{aligned} h_D &= c\underline{\alpha}\cdot(\underline{p} - e\underline{A}) + \beta mc^2 + e\Phi \\ &= \beta mc^2 + \mathcal{O} + \mathcal{E} \end{aligned} \tag{9.1}$$

(with Φ the Coulomb field) where the "odd", $\mathcal{O}$, and "even", $\mathcal{E}$, parts satisfy

$$\begin{aligned} \mathcal{O} &= c\underline{\alpha}\cdot(\underline{p} - e\underline{A}), & \{\beta,\mathcal{O}\} &= 0, \\ \mathcal{E} &= e\Phi, & [\beta,\mathcal{E}] &= 0 \end{aligned} \tag{9.2}$$

Any operator H can be separated to

$$\begin{aligned} \mathcal{O} &= \frac{1}{2}(H - \beta H\beta) \\ \mathcal{E} &= \frac{1}{2}(H + \beta H\beta) \end{aligned} \tag{9.3}$$

We shall annihilate the "odd" part up to some power of $1/mc^2$ through the transformation

$$\Psi' = \mathrm{e}^{\mathrm{i}S}\Psi \tag{9.4}$$

A good choice is

$$S = \frac{-\mathrm{i}\beta\mathcal{O}}{2mc^2} = -\mathrm{i}\beta\frac{\underline{\alpha}\cdot\underline{\pi}}{2mc} \tag{9.5}$$

where $\underline{\pi}$ is represents the generalized momentum.

A third order FW Hamiltonian becomes

$$
\begin{aligned}
h_D''' =& mc^2 ? e\Psi + \frac{1}{2m}\pi^2 && && (9.6a)\\
&- \frac{e}{m}\beta \underline{s}\cdot\underline{B} && && (9.6b)\\
&- \frac{1}{8m^3c^2}p^4 && \text{(mass-velocity)} && (9.6c)\\
&+ \frac{e}{8m^2c^2}\nabla\cdot\underline{E} && \text{(Darwin)} && (9.6d)\\
&- \frac{e}{2m^3c^2}\underline{S}\cdot\underline{E}\times\underline{p} && \text{(spin-orbit)} && (9.6e)\\
&- \frac{e\hbar}{m}\frac{1}{2m^2c^2}(\underline{S}\cdot\underline{B})\underline{\pi}^2 && \text{(rel. corr. to Zeeman)} && (9.6f)
\end{aligned}
$$

e is the charge of the particele, $\hbar = 1$, $s = \frac{1}{2}\sigma$. [See **Bjorken-Drell** (1964), eq. (4.5), **Moss** (1973), eq. (9.75)].

This expression can be used in 1st-order perturbation theory like the Pauli one. It is not useful in variational calculations as both the mass-velocity term $h_m = -\frac{1}{8c^2}p^4$ and the Darwin term $h_D = \frac{1}{8c^2}\nabla\underline{E} \overset{V_n}{=} -\frac{Ze\pi}{2c^2}\delta(\underset{\sim}{r})$ are strongly singular. One could just as well use them to first order, $\mathcal{O}(c^{-2})$. Contracted basis sets could however allow variational use.

9.3 The Cowan-Griffin equation

Cowan and **Griffin** (1976) or **Wood** and **Boring** (1978) as well as other authors use the 2nd-order equation of Dirac (1928a), putting all the norm in the large component.

$$
\begin{cases}
P' + \frac{\kappa}{r}P + (-mc + \frac{V-E}{c})Q = 0\\
Q' - \frac{\kappa}{r}Q + (-mc + \frac{E-V}{c})P = 0
\end{cases}
\tag{9.7}
$$

$$
Q = \frac{c}{mc^2 + E - V}(P' + \frac{\kappa}{r}P) \tag{9.8}
$$

$$
Q' = \frac{c}{mc^2 + E - V}\big(P'' + \frac{\kappa}{r}P' - \frac{\kappa}{r^2}P\big) + \frac{cV'}{(mc^2 + E - V)^2}\big(P' + \frac{\kappa}{r}\big)P \tag{9.9}
$$

$$\frac{c}{mc^2+E-V}\big(P''+\frac{\kappa}{r}P'-\frac{\kappa}{r^2}P\big)-\frac{\kappa}{r}\frac{c}{mc^2+E-V}\big(P'+\frac{\kappa}{r}P\big)$$
$$+\frac{cV'}{(mc^2+E-V)^2}\big(P'+\frac{\kappa}{r}P\big)+\frac{-mc^2+E-V}{c}P=0 \quad |E=mc^2+\varepsilon$$
$$P''+\underline{\frac{\kappa}{r}P'}-\frac{\kappa}{r^2}P-\underline{\frac{\kappa}{r}P'}-\frac{\kappa^2}{r^2}P+\frac{V'}{2mc^2+\varepsilon-V}\big(P'+\frac{\kappa}{r}P\big)$$
$$+\frac{1}{c^2}\big(2mc^2+\varepsilon-V\big)\big(\varepsilon-V\big)P=0$$
$$P''-\frac{\kappa(\kappa+1)}{r^2}P+2m(\varepsilon-V)P\underbrace{+\frac{V'}{2mc^2+\varepsilon-V}\big(P'+\frac{\kappa}{r}P\big)}_{h_{\rm d}+h_{\rm SO}}$$
$$+\underbrace{\frac{1}{c^2}(\varepsilon-V)^2}_{h_m}=0 \tag{9.10}$$

$$\kappa(\kappa+1)=l(l+1) \tag{9.11}$$

Quasirelativistic usage: put $\int_0^\infty P^2{\rm d}r=1$! (This is the only approximation!)

Dirac (1928) [PRSL A **117** (1928) 610]: Exact 2nd-order equation.

For the **non-relativistic** case,

$$r\to 0: \quad P''=\frac{l(l+1)}{r^2}P \quad \Rightarrow P\propto r^{l+1}, \quad R\propto\frac{P}{r}=r^l \tag{9.12}$$
$$r\to\infty: \quad P''=-2m\varepsilon P \quad \Rightarrow P\propto {\rm e}^{\xi r}, \quad \xi=\sqrt{-2m\varepsilon} \tag{9.13}$$

For the **relativistic** case,

$$p\propto r^\gamma, \quad \gamma=\sqrt{\kappa^2-(\frac{z}{c})^2} \tag{9.14}$$

9.4 Douglas-Kroll-Hess

Instead of expanding in powers of $1/c$ one can rather expand in powers of the coupling strength, $Z\alpha c\hbar$ (**Douglas** and **Kroll** 1974, Hess 1986, **Jansen** and Hess 1989). This leads to operators which are bound from below and can be used variationally.

The DK formalism, which is an *all electron* method, is based on a series of unitary transformations $U_0, U_1, \ldots$ of which the lowest is the *free-particle FW-transformation* defined by

$$\begin{cases} U_0=A(1+\beta R) \\ U_0^{-1}=(R\beta+1)A \end{cases} \tag{9.15}$$

where we have

$$A = \sqrt{\frac{E_p + mc^2}{2E_p}} \tag{9.16}$$

$$R = \frac{c\underline{\alpha} \cdot \underline{p}}{E_p + mc^2} \tag{9.17}$$

$$E_p = c\sqrt{p^2 + m^2c^2} \tag{9.18}$$

Applying U_0 to $h_D = c\underline{\alpha} \cdot \underline{p} + (\beta - 1)mc^2 + V$ gives

$$U_0 h_D U_0^{-1} = \beta E_p + \mathcal{E}_1 + \mathcal{O}_1 \equiv H_1 \tag{9.19}$$

with

$$\begin{aligned} \mathcal{E}_1 &= A(V + RVR)A \\ \mathcal{O}_1 &= \beta A(RV - VR)A \end{aligned} \tag{9.20}$$

For chemical purposes one more transformation is needed. One uses

$$U_1 = \sqrt{1 + W_1^2} + W_1 \tag{9.21}$$

with W_1 anti-hermitian, $W_1^\dagger = -W_1$. Performing the transformation through U_1 and expanding the square root in powers of W_1,

$$\begin{aligned} U_1 H_1 U_1^{-1} = \beta E_p &- [\beta E_p, W_1] + \mathcal{E}_1 + \mathcal{O}_1 \\ &+ \frac{1}{2}\beta E_p W_1^2 + \frac{1}{2}W_1^2 \beta E_p - W_1 \beta E_p W_1 \\ &+ [W_1, \mathcal{O}_1] + [W_1, \mathcal{E}_1] + \ldots \end{aligned} \tag{9.22}$$

omitting higher order terms. The first-order odd term is eliminated by setting

$$[\beta E_p, W_1] = \mathcal{O}_1 \tag{9.23}$$

and solving for W_1. It's momentum space integral operator is

$$W_1 \Phi(\underline{p}) = \int d^3p' W_1(\underline{p}, \underline{p}')\Phi(\underline{p}') \tag{9.24}$$

with a kernel

$$W_1(\underline{p}, \underline{p}') = A(R - R')A' \frac{V(\underline{p}, \underline{p}')}{E_{p'} + E_p} \tag{9.25}$$

$V(\underline{p}, \underline{p}')$ is the Fourier transform of the external potential. The primed quantities are expressed in terms of the variable $\underline{p}'$. The final result is

$$H^{\text{decoupled}} \cong \beta E_p + \mathcal{E}_1 - \beta\big[W_1 E_p W_1 + \frac{1}{2}[W_1^2, E_p]\big]. \tag{9.26}$$

The spin-orbit terms can be separated using

$$(\underline{\sigma}\cdot\underline{A})(\underline{\sigma}\cdot\underline{B}) = \underline{A}\cdot\underline{B} + \mathrm{i}\underline{\sigma}\cdot(\underline{A}\times\underline{B}) \tag{9.27}$$

see Hess *et al.* (1995). The final result becomes

$$\begin{aligned} H_{\mathrm{SO}}^{+} = &\sum_{i}^{\text{el.}}\sum_{\alpha}^{\text{nuc.}} e^2 Z_\alpha \frac{A_i}{E_i+m}\underline{\sigma}_i\cdot\Big(\frac{\underline{r}_{i\alpha}}{r_{i\alpha}}\times\underline{p}_i\Big)\frac{A_i}{E_i+m} \\ &- e^2\sum_{i\neq j}\frac{A_iA_j}{E_i+m}\Big(\frac{\underline{r}_{ij}}{r_{ij}^3}\times\underline{p}_i\Big)\cdot(\underline{\sigma}_i+2\underline{\sigma}_j)\frac{A_iA_j}{E_i+m} \end{aligned} \tag{9.28}$$

A series expansion in powers of c^{-2} becomes

$$\frac{A_i}{E_i+mc^2} = \frac{1}{2mc^2} - \frac{3p_i^2}{16m^3c^4} + \ldots \tag{9.29}$$

The DK Hamiltonian has now been implemented in a large number of programs, to mention a few: MAGIC, MOLCAS, PARAGAUSS and those by Hess himself.

For calculating properties, the operator has to be transformed as well, if transformed Hamiltonians are used. This could be quite tedious.

9.5 Zero Order Regular Approximation, ZORA

The previous elimination procedures treated $\frac{E-V}{mc^2}$ as a small parameter. Sufficiently near the nucleus this is not true. Here

$$\omega(\underline{r}) = \Big(1-\frac{V-E}{2mc^2}^{-1}\Big) = \sum_{n=0}^{\infty}\Big(\frac{V-E}{2mc^2}\Big)^n \tag{9.30}$$

will only converge when $|\frac{V-E}{2mc^2}|<1$. By rewriting the expression as

$$\omega(\underline{r}) = \frac{2mc^2}{2mc^2-V}\Big(1+\frac{E}{2mc^2-V}\Big)^{-1} \tag{9.31}$$

we obtain an expression in

$$\frac{1}{2mc^2-V}$$

which will converge. This leads to the Zero (or higher) Order Regular Approximation, **ZORA** (**E. van Lenthe** *et al.* 1993–).

The ZORA-level approximation was already given by **Chang**, **Pélissier** and **Durand** (1986) or by **Heully** *et al.* (1986).

One problem with ZORA is the electric gauge invariance but this can be approximately remedied (E. van Lenthe, **Baerends**, **Snijders** 1994, **van Wüllen** 1998).

The scalar scaled ZORA equation is (**Faas** *et al.* 2000)

$$\frac{1}{S}\left[\underline{p}\cdot B\underline{p}+V_c-K_{\Psi\Psi}\right]\psi_i=\mathcal{E}_i^{\text{scaled}}\psi_i \tag{9.32}$$

with the scaling factor

$$S=1+\langle\Psi_i|\underline{p}\cdot\frac{1}{c^2}B^2\underline{p}|\Psi_i\rangle \tag{9.33}$$

$$B=\frac{c^2}{2c^2-V_c} \tag{9.34}$$

$$V_c=V_{\text{nuc}}+J_{\psi\psi}$$

J and K are the usual Coulomb and exchange operators, respectively. Van Wüllen's idea was to set $B=B_0=c^2/(2c^2-V_0)$, V_0 depending on the atoms only.

ZORA has been implemented in packages as GAMESS-UK, ADF and DIRAC.

F. Wang *et al.* (2000) include spin-orbit effects at DFT level in

$$\left[\underline{\sigma}\cdot\underline{p}\frac{c^2}{2c^2-V_0}\underline{\sigma}\cdot\underline{p}+V\right]\phi=\mathcal{E}\phi \tag{9.35}$$

The separated SO operator becomes

$$h_{\text{SO}}=\mathrm{i}\underline{\sigma}\cdot\left[\underline{p}\times\frac{V_0}{2(2c^2-V_0)}\underline{p}\right] \tag{9.36}$$

9.6 Direct Perturbation Theory, DPT

The wave function and its metric will have to be an analytic function at the limit $c^{-1}\to 0$.

The Direct Perturbation Theory (DPT) has been developed by **Sewell** (1949), **Titchmarsh** (1962), **Rutkowski** (1986–), **Kutzelnigg** (1989–). See also **Franke** (1994).

Instead of the original Dirac equation

$$\begin{pmatrix}V-E & c\underline{\sigma}\cdot\underline{p}\\ c\underline{\sigma}\cdot\underline{p} & -2mc^2+V-E\end{pmatrix}\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix}=0 \tag{9.37}$$

one now writes

$$\begin{pmatrix}V-E & \underline{\sigma}\cdot\underline{p}\\ \underline{\sigma}\cdot\underline{p} & -2m+t(V-E)/c^2\end{pmatrix}\begin{pmatrix}\psi_+\\ c\psi_-\end{pmatrix}=0 \tag{9.38}$$

with the parameter $t = 0$ and $t = 1$ corresponding to the non-relativistic and relativistic cases respectively. The NR case corresponds to the Lévy-Leblond (1967) equation. Both E and

$$\Psi = \begin{pmatrix} \psi_+ \\ c\psi_- \end{pmatrix}$$

are expanded in powers of t (or c^{-2}). A key paper on the subject is Kutzelnigg (1996).

9.7 Further examples

9.7.1 RESC

Hirao (see **Nakajima** *et al.* 1999, 2000) start from the inconvenience of having the energy in the 2nd-order equation

$$\left[V + (\underline{\sigma} \cdot \underline{p}) \frac{c^2}{2mc^2 - V + E} (\underline{\sigma} \cdot \underline{p})\right] \psi_{\mathrm{L}} = E\psi_{\mathrm{L}} \tag{9.39}$$

$$\langle \psi_{\mathrm{L}} | 1 + x^\dagger x | \psi_{\mathrm{L}} \rangle = 1, \quad \psi_{\mathrm{S}} = X\psi_{\mathrm{L}}$$

They suggest replacing

$$E - V \rightarrow T = \sqrt{m^2c^4 + p^2c^2} - mc^2 \tag{9.40}$$

which is independent of the eigenvalue E. The Hamiltonian can again be separated into spin-free and spin-dependent parts.

Barysz (2000) has analyzed this family of **RESC** (Relativistic scheme for Eliminating Small Components) Hamiltonians. Although they can show signs of variational collapse already at moderate Z (≈ 80), the result for valence electrons can be close to Douglas-Kroll ones.

9.7.2 AMFI

The Atomic-Mean-Field-Integral method eliminates the spin-other-orbit terms in terms of an effective one-electron, one-centre potentials. (Hess *et al.* 1996, Marian and **Wahlgren** 1996, Schimmelpfennig *et al.* 1998ab, **Tatchen** and Marian 1999).

An average over the two-electron terms is taken for the valence shell, and thus large two-electron integral files can be avoided.