Lecture I: Relativistic effects, energy levels in free metal-ions: multiplets, spin-orbit coupling and fine structure, magnetic effects in atomic structure.

Relativistic Effects in Chemistry: a road map

Example: Are atoms relativistic?

The relativistic effect are important whenever the speed $v$ is comparable to the speed of light $c$ ($c=2.99792458 \times 10^8 \text{m/s}$). In atomic units, the speed of light is $c=137.0359895$. For "non-relativistic atoms", i.e. the atoms described by the non-relativistic Schrödinger equation, the average speed of a 1s electron is equal to $Z$ (in atomic units). Therefore,

$$\frac{v_{av}(1s)}{c} = \frac{Z}{137}$$

For hydrogen, this ratio is small $1/137$ but for the atoms like Au ($Z=79$) or Hg ($Z=80$) it is about 58%.

We recall now the formula for the relativistic mass increase:

$$m = \frac{m_0}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

and the definition of the Bohr radius:

$$a_0 = \left(4\pi\varepsilon_0\right)\left(\frac{\hbar^2}{me^2}\right)$$

The radius of a 1s orbital in the relativistic case is smaller than its non-relativistic counterpart. For gold the radial shrinkage of the 1s orbital amounts to 23%.

Example: Exact solution for the hydrogen-like $Z=80$ cation?

For Au$^{79}$, the exact non-relativistic and the relativistic solutions can be obtained (cf. Appendix 1 and [Proc. Phys. Soc. London, 90 (1967) 297]).

The relativistic level of description is needed for many problems in chemistry of heavy elements. The "direct relativistic effects" include:

Spin-orbit coupling. Important for all magnetic and magnetic resonance properties of metal complexes!
**Shrinking of s and p orbitals.** For instance, it amounts to about 20% in mercury. Shrinkage of the core orbitals leads to a number of important chemical consequences. The electronic charge of the nucleus is better screened by shrunk orbitals which leads to expansion of chemically active outer shell d and f valence orbitals.

Some examples (for more, see Pekka Pykkö [Chem. Rev., 88 (1998) 563]):

- **Yellow color of gold** (solid). The adsorption at 2.4 eV is assigned to the transition from filled 5d band to the Fermi level (6s). The rising of the 5d level is due to the relativistic effect. For silver, the rising is smaller and the transition occurs at 3.7 eV (ultraviolet).

- **Electron affinity of gold.** Relativistic effects allow to form an ionic Au$^-$ ion in CsAu and RbAu crystals (semiconductors).

- **Contraction of the bond lengths** (for PbH$_4$ and AuH amounting to 6-7%).

- **Lanthanide Contraction:** along the series from La(III) (4f$^0$) to Lu(III) (4f$^{14}$) the ionic radii decrease by 18.3 pm. About 10% of this effects has a relativistic origin.

**Can the Schrödinger equation be applied for relativistic molecules?**

Unfortunately **no**. The kinetic energy operator in the Schrödinger equation corresponds to the kinetic energy of a non-relativistic particle and there is nothing in the Schrödinger equation which would prohibit the value of averaged velocity to reach c or even to exceed it.

Moreover, the Schrödinger equation for an electron:

\[
i \frac{\partial \Psi}{\partial t} = \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + V(x,y,z) \right] \Psi = \hat{H} \Psi
\]

cannot be reconciled with the special theory of relativity in which the space (x, y, z) and time (t) coordinates are treated on equal footing (Lorentz transformation).

In 1928, Dirac suggested an other equation (cf. Appendix 1) which makes the quantum theory compatible with relativity:

\[
-i \hbar \left( \hat{\alpha}_x \frac{\partial}{\partial x} + \hat{\alpha}_y \frac{\partial}{\partial y} + \hat{\alpha}_z \frac{\partial}{\partial z} \right) + \beta mc^2 + V(x,y,z) \right] \Psi = i \frac{\partial \Psi}{\partial t}
\]

In the above equations $\hat{\alpha}$ and $\hat{\beta}$ are not scalars but they are 4x4 matrices of known coefficients. The relativistic wavefunction contains, therefore, four components:

\[
\Psi = \begin{pmatrix}
\psi_1 \\
\psi_2 \\
\psi_3 \\
\psi_4
\end{pmatrix}
\]
Both the Dirac and the Schrödinger equations cannot be solved exactly for many electron systems. Similarly as in the case of the Schrödinger equation, many *approximate* methods to solve the Dirac equation have been developed. These methods will be reviewed here. The four components $\Psi_1, \Psi_2, \Psi_3, \Psi_4$ of the relativistic wavefunction $\Psi$ are written frequently as:

$$\Psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}$$

where each $\phi$ and $\chi$ comprise only two components. The explicit form of the time-independent Dirac equation $\phi$ and $\chi$ takes the following form:

$$V\phi + c\vec{\sigma} \cdot \vec{p}\chi = E\phi$$
$$c\vec{\sigma} \cdot \vec{p}\phi + (\vec{V} - 2c^2)\chi = E\chi$$

where traditionally $\phi$ and $\chi$ are called *large component* and *small component*, respectively. The $2\times2$ matrices $\vec{\sigma}$ are Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$
$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Elimination of the small component from the Dirac equation leads to the following equation for the large component $\phi$:

$$\hat{H}^{esc}\phi = V\phi + \frac{1}{2}\vec{\sigma} \cdot \vec{p} \left( 1 + \frac{E - V}{2c^2} \right)^{-i} \vec{\sigma} \cdot \vec{p}\phi = E\phi \quad (*)$$

where $\hat{H}^{esc}$ denotes the Dirac Hamiltonian with *eliminated small component*.

Equation (*) provides a particular form of the Dirac equation which is useful in further discussions. Comparing it to the non-relativistic Schrödinger equation shows that the kinetic energy operator is replaced by a more complicated object:

$$\frac{1}{2}\vec{p}^2 \to \frac{1}{2}\vec{\sigma} \cdot \vec{p} \left( 1 + \frac{E - V}{2c^2} \right)^{-i} \vec{\sigma} \cdot \vec{p}$$

Solving Dirac equation is a more complicated task than solving the corresponding Schrödinger equation. Note that $E$ appears in the left-hand-side and in the right-hand-side of the equation (*).

The figs. next show radial wavefunctions of heavy noble gases.
Approximate solutions of Dirac equation: Pauli Hamiltonian

Expanding the $\hat{H}^{exc}$ Hamiltonian in $\frac{P^2}{c^2}$ and keeping the zeroth and first order terms leads to the approximate relativistic Hamiltonian known as Pauli Hamiltonian:

$$H^{Pauli} = V + \frac{P^2}{2} - \frac{P^4}{8c^2} + \frac{\nabla^2 V}{8c^2} + \frac{1}{4c^2} \hat{\sigma} \cdot (\nabla \times \mathbf{p})$$

where:

- $-\frac{P^4}{8c^2}$ is called mass-velocity term;
- $\frac{\nabla^2 V}{8c^2}$ is called Darwin term;
- $\frac{1}{4c^2} \hat{\sigma} \cdot (\nabla \times \mathbf{p})$ is called spin-orbit term.

Neglecting the last term in the above equation leads to so called scalar relativistic Pauli Hamiltonian.

This method has been implemented in the ADF program.

Different wave-function based approaches have been combined with relativistic calculations based on Dirac’s equation (at different level of approximation to the exact Dirac Hamiltonian). The wave-function approaches will not be discussed here. It is worthwhile to know that also the X-α model which is a prototype of modern DFT methods was applied for relativistic calculations (see the review article by Pekka Pykkö [Chem. Rev., 88 (1998) 563]).

The four-component Dirac-Kohn-Sham equations take the following form:

$$\left\{-i\hat{\alpha} \cdot \nabla - m(1 - \hat{\beta}) + V_{\text{eff}}[\mathbf{r}; j_\mu]\right\} \Psi^{KS}(\mathbf{r}) = \varepsilon_i \Psi^{KS}(\mathbf{r})$$

where:

$$V_{\text{eff}}[\mathbf{r}; j_\mu] = -\left(\nabla \varepsilon_{\text{ext}}(\mathbf{r}) + \int \frac{\mathbf{r}' \cdot \mathbf{j}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r}' + \frac{\delta E_{\text{XC}}}{\delta \rho(\mathbf{r})}\right)$$

The ground-state electron $\rho(\mathbf{r})$ and current densities $\mathbf{j}(\mathbf{r})$ are given by:

$$\rho(\mathbf{r}) = \sum_i N_{\text{occ}} \text{Tr}\{\Psi^{KS*}_i(\mathbf{r}) \Psi^{KS}_i(\mathbf{r})\}$$
and

\[ j_i(r) = \sum_j^{N_{\text{occ}}} \text{Tr} \{ \Psi_j^{KS}(r) \bar{\alpha}_i \Psi_i^{KS}(r) \} \]

where \( \Psi_i^{KS}(r) \) denote the relativistic Dirac-Kohn-Sham orbitals.

Note that, the relativistic exchange-correlation potential depends not only on the electron and spin densities, but it depends on four current densities (electron density, spin-density, and orbital current density in the non-relativistic limit) [Rajagopal and Callaway, Phys. Rev. B, 7 (1973) 1912].

The good approximation for the relativistic exchange-correlation can be obtained by considering the non-relativistic exchange-correlation functional as a leading term and adding corrections due to magnetic interactions and retardation (Breit interactions) [Rajagopal and Ramana, Adv. Chem. Phys., 54 (1983) 231]. The leading term has the non-relativistic form:

\[ E_{\text{xc}}[\rho, s] = \int \rho \epsilon_{\text{xc}}(\rho, s) dr \]

where the spin density is obtained from the relativistic four-component wavefunction:

\[ s = \sum_i^{N_{\text{occ}}} \phi_i \bar{\sigma} \phi_i \sum_i^{N_{\text{occ}}} \chi_i \bar{\sigma} \chi_i \]

where \( \phi_i \) and \( \chi_i \) are the large and the small component of the relativistic four component wavefunction \( \Psi \).

**Approximate solutions of Dirac equation: CPD Hamiltonian (ZORA)**

Chang, Pélisier, and Durand [Phys.Scr., 34 (1986) 394] derived another expansion of the \( \hat{H}^{\text{exc}} \). In its zero-th order, the Hamiltonian takes the form:

\[ H^0 = V + \bar{\sigma} \cdot \mathbf{p} \frac{e^2}{2c^2 - V} \bar{\sigma} \cdot \mathbf{p} \]

\[ = V + \mathbf{p} \frac{e^2}{2c^2 - V} \mathbf{p} + \frac{e^2}{2c^2 - V} \bar{\sigma} \cdot (\nabla V \times \mathbf{p}) \]

The same zero-th order term has been derived by van Lenthe, Baerends, and Snijders [J.Chem.Phys., 99 (1993) 4597]. It is known as ZORA (Zeroth Order Regular Approximation) approach. This approach is currently implemented in the ADF program. The last term in the above equation is the spin-orbit term. Neglecting it leads to a scalar equation (scalar-ZORA). The following paper illustrates the significance of this term in closed-shell relativistic molecules [J. Chem.Phys., 105 (1996) 6505].

**Relativistic Effective Core Potentials (ECP)**

Exact and approximate solutions of the Dirac equation show invariably that the s-orbitals of relativistic atoms shrink and their orbital energies are lowered compared to their hypothetic non-relativistic counterparts.
We recall now the concept of pseudopotentials discussed elsewhere [Phys. Rev., 116 (1959) 287]. Pseudopotentials make it possible to separate the core and the valence electrons. One can use the relativistic calculations to obtain s orbitals of the core and then apply usual non-relativistic Schrodinger equation in its pseudopotential formulation for remaining electrons. See for example the work of Kahn, Hay and Cowan [J. Chem. Phys., 68 (1978) 2386]. The concept of pseudopotentials is, therefore, very useful for relativistic elements. The most 'relativistic' electrons are removed from the problem.

The relativistic effective core potentials are also applied within the Kohn-Sham formalism. According to Prof. Pekka Pykkö, in 90% of cases relativistic ECP provide an adequate theoretical approach to obtain the geometry and electronic structure of complexes involving relativistic elements.

FIG. 6. Valence wave functions for gold. Dashed curves correspond to the Dirac large component for the full-core atom, the solid curves to the pseudopotential (see Fig. 5) atom.

**Magnetic effects in atomic structure**

**Interaction with an external magnetic field : ZEEMAN Effect.**

We consider here the electronic Zeeman effect. However, since many nuclides have non-zero *intrinsic* spin, the same interaction occurs with an external field. The magnitude of this interaction is, however, several (3-4) orders of magnitude smaller.

Hamiltonian: 

\[ \hat{H}_{Ze} = \vec{\mu} \cdot \vec{H} \]

\( \vec{\mu} : \) magnetic dipole moment  
\( \vec{H} : \) magnetic field

**Contribution to the angular momentum:**

\[ \vec{\mu}_L = \vec{I} \cdot S \]

**current:**  
\[ \vec{I} = -e \vec{\omega} = -e \frac{\vec{v}}{2\pi r} \]

**surface:**  
\[ S = \pi r^2 \]

\[ \vec{\mu}_L = \vec{I} \cdot S = \left( -e \frac{\vec{v}}{2\pi r} \right) \cdot \left( \pi r^2 \right) = \frac{-e\vec{v}r}{2c} \text{ [emu]} \]

But  
\[ \vec{p}_L = m\vec{v}r \]

thus  
\[ \vec{\mu}_L = \left( -\frac{e}{2mc} \right) \vec{p}_L \]

in quantum mechanics:  
\[ \vec{p}_L \leftrightarrow -i\hbar \hat{L} \]

thus  
\[ \vec{\mu}_L = \frac{e\hbar}{2mc} \hat{L} = \beta_e \hat{L} \]
Spin contribution

\[ \vec{\mu}_S = g_e \beta_e \vec{S} \]

ge = 2.0023  gyromagnetic factor of electron

Thus, the Zeeman hamiltonian \( \hat{H}_{Ze} \) reads:

\[ \hat{H}_{Ze} = \beta_e (\hat{L} + g_e \vec{S}) \cdot \hat{H} \]

For nuclei with a nuclear spin, the complete Zeeman hamiltonian \( \hat{H}_{Ze}^{tot} \) is thus:

\[ \hat{H}_{Ze}^{tot} = \beta_e (\hat{L} + g_e \vec{S}) \cdot \hat{H} - g_N \beta_N \hat{I} \cdot \hat{H} \]

Where \( g_N \) and \( \beta_N = \frac{e\hbar}{2Mc} \) are respectively the gyromagnetic factor and the Bohr magneton of the nucleus.

Paramagnetic susceptibility: Van Vleck’s formula

This formula describes the dependence of the magnetic susceptibility \( \chi(T) \), i.e. the proportionality factor between the bulk magnetisation and the external magnetic field, as a function of temperature and is the basis for the whole magnetochemistry.

Consider the energy \( E_n \) of the n-th electronic level of an ion or complex molecule in an external magnetic field \( \vec{H} \) and suppose that \( E_n \) can be expanded in a Taylor series as:

\[ E_n = E_n^{(0)} + E_n^{(1)}H + E_n^{(2)}H^2 + ... \]

where \( E_n^{(0)} \) is the energy in absence of a magnetic field and where \( E_n^{(1)} \) and \( E_n^{(2)} \) the 1st and 2nd order Zeeman coefficients respectively. To each level \( n \), Van Vleck associates a magnetic moment \( \mu_n \), defined as:

\[ \mu_n = \frac{\partial E_n}{\partial H} \]

The total magnetic moment \( M \), i.e. the bulk magnetisation, is obtained in supposing that the levels \( n \) are populated according to a Boltzmann distribution, that is:
\[
M = \frac{N \sum_n \mu_n \exp\left(-\frac{E_n}{kT}\right)}{\sum_n \exp\left(-\frac{E_n}{kT}\right)}
\]

where \(k\) denotes the Boltzmann constant, \(T\) the temperature and \(N\) the Avogadro number.

If \(H\) is sufficiently small, we can approximate:

\[
\exp\left(-\frac{E_n}{kT}\right) \approx \exp\left(-\frac{E_n^{(0)}}{kT}\right) \left[1 - \frac{E_n^{(1)}H}{kT}\right]
\]

\(M\) reads then:

\[
M = \frac{N \sum_n \left(-E_n^{(1)} - 2E_n^{(2)}H\right) \left[1 - \frac{E_n^{(1)}H}{kT}\right] \exp\left(-\frac{E_n^{(0)}}{kT}\right)}{\sum_n \left[1 - \frac{E_n^{(1)}H}{kT}\right] \exp\left(-\frac{E_n^{(0)}}{kT}\right)}
\]

For \(H = 0\), \(M\) must vanish when the molecule or atom has no permanent magnetic moment, thus:

\[
\sum_n E_n^{(1)} \exp\left(-\frac{E_n^{(0)}}{kT}\right) = 0
\]

Keeping in the expression for \(M\) only terms linear in \(H\), we can express the paramagnetic susceptibility as:

\[
\chi = \frac{M}{H}
\]

which is nothing but Van Vleck formula:

\[
\chi = \frac{N \sum_n \left[\frac{(E_n^{(1)})^2}{kT} - 2E_n^{(2)}\right] \exp\left(-\frac{E_n^{(0)}}{kT}\right)}{\sum_n \exp\left(-\frac{E_n^{(0)}}{kT}\right)}
\]

describing the temperature dependence of \(\chi\) as a function of temperature. Due to the approximation of the exponential function for small arguments, the formula above might lose its validity say below 10K.

The coefficients \(E_n^{(1)}\) and \(E_n^{(2)}\) can easily be obtained from perturbation theory. Indeed, let \(\ket{n}\) be the wave function associated to level \(n\) and adapted to the perturbation \(\hat{\mu} \cdot \vec{H}\), we get:

\[
E_n^{(1)} = \langle n | \hat{\mu} | n \rangle
\]

and

\[
E_n^{(2)} = \sum_{m 
eq n} \frac{\langle m | \hat{\mu} | n \rangle^2}{E_n^{(0)} - E_m^{(0)}}
\]
the summation in the latter expression is restricted to indices \( m \) such that \( E_n^{(0)} \neq E_m^{(0)} \).

**Example**: Paramagnetic susceptibility of rare earth ions.

⇒ magnetism is essentially **independent of environment**

- Magnetic moment of a \( J \)-state is expressed by **Landé’s formula**

\[
\mu = g_J \sqrt{J(J+1)} \mu_B; \quad g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)}
\]

---

**Sample Landé Calculation**

*E.g Pr\(^{3+}\) [Xe]4f\(^2\)*

- Find Ground State from Hund's Rules
  - **Maximum Multiplicity** \( S = \frac{1}{2} + \frac{1}{2} = 1 \Rightarrow 2S + 1 = 3 \)
  - **Possible Orbital Angular Momenta** \( L = 0 - 6 \Rightarrow S, [P], [D], [F], [G], [H], [I] \)
    where anti-symetrized squares are in [brackets].
  - **Hund’s rule**: (i) max. \( S \) and (ii) max. \( L \) \( \Rightarrow ^3H \) state
  - **Total Angular Momentum** \( J = (L + S) \), \((L + S) - 1\), ..., \(|L - S| = 6\), 5, 4
    - Less than half-filled sub-shell \( \Rightarrow \) Minimum \( J \) \( \Rightarrow J = 4 \)
      - \{Greater than half-filled sub-shell \( \Rightarrow \) Maximum \( J \}\}
  - **Experiment? 3.4 - 3.6 \( \mu_B \)**

---

Ln\(^{3+}\) Magnetic Moments compared with Theory
Exercice: Repeat this example using the methods of chap. II.

Kramers degeneracy
In the absence of an external magnetic field, molecules with unpaired electrons (odd number of electrons) are at least two-fold degenerate. This is a fundamental theorem of quantum chemistry which traces back to time reversal symmetry.
Appendix 1: The Dirac Equation

One of the first attempts at a relativistic quantum mechanical wave equation was the Klein-Gordon equation. Though it is theoretically sound from the perspective that it is consistent with both classical quantum mechanics and the special theory of relativity, it has several unsavory features which keep it from being a very powerful tool in relativistic quantum mechanics. Dirac later developed his own relativistic wave equation which did not have some of the shortcomings of the Klein-Gordon equation and that some of the supposed “errors” that the KG equation gave rise to were actually illustrating some new physics. The Dirac equation is only rigorous for a one particle system, but has been used as a starting point for a number of approximate many electron methods.

1.1 Klein-Gordon Equation
The special-relativistic expression for the kinetic energy may be used to form the classical free particle Hamiltonian

\[ E = c \left( m^2 c^2 + p^2 \right)^{1/2} \]  (1.1)

The analogous quantum mechanical expression may be constructed by replacing the classical momentum, \( p \), with its quantum mechanical operator, \( \hat{p} \) which yields the free particle wave equation

\[ \left( i\hbar \frac{\partial}{\partial t} \right) \psi = \left[ c \left( m^2 c^2 + \hat{p}^2 \right)^{1/2} \right] \psi \]  (1.2)

This equation, however, does not satisfy some of the conditions required by special relativity. The wave equation is not invariant to a Lorentz transform, and the square root term introduces ambiguity. The Klein-Gordon equation, first proposed in 1927 rectifies both of these problems simply by taking the square of the original energy expression and extending the result to a quantum mechanical wave equation:

\[ E^2 \psi = \left( i\hbar \frac{\partial}{\partial t} \right)^2 \psi = \left( m^2 c^4 + \hat{p}^2 c^2 \right) \psi \]  (1.3)

The resulting wave equation is Lorentz invariant and well defined, but suffers from other problems. Negative energy solutions to this equation are possible, which do not have a readily obvious explanation, and the probability density, \( \psi^* \psi \), fluctuates with time as does its integral over all space.

The foibles of the Klein-Gordon equation may make it a poor equation for the electron, but those weaknesses helped to point Dirac in the right direction and to develop a single particle equation which successfully surmounted all these problems.

1.2 Dirac’s Free Particle Equation
In order for a wave equation to satisfy the special relativistic requirement of Lorentz invariance, derivatives in space and time must all appear in the same order. The K-G equation illustrates that an expression which satisfies this condition but is non-linear in the space and time derivatives gives rise to anomalous results. Dirac set out to find an equation which was first order in space and time derivatives. The result of his efforts, the Dirac equation, is difficult to motivate, impossible to prove, and far more complicated than the non-relativistic analog. However, Dirac’s wave equation for a single particle satisfies all the requirements of special relativity and quantum mechanics, and is able to predict the properties of one particle systems with remarkable accuracy.
Dirac’s equation for an electron in field-free space is given by

\[
(\hat{p}_0 - \alpha \cdot \hat{p} - \beta mc)\psi = 0 \tag{1.4}
\]

where

\[
\hat{p}_0 = \frac{i}{\hbar} \frac{\partial}{\partial t}
\tag{1.5}
\]

and \(\hat{p}\) is simply the three component momentum. In order to determine the nature of the three components of \(\alpha\) and \(\beta\), it is useful to compare the modified equation

\[
(\hat{p}_0 - \alpha \cdot \hat{p} - \beta mc)^*(\hat{p}_0 - \alpha \cdot \hat{p} - \beta mc)\psi = 0 \tag{1.6}
\]

to the Klein-Gordon equation. Equations (1.3) and (1.6) are equivalent if we enforce the conditions

\[
[\alpha_i, \alpha_j] = \alpha_i \alpha_j + \alpha_j \alpha_i = \delta_{ij} \tag{1.7}
\]

where \(\alpha_i\) represents \(\beta\) for \(i=0\) and \(\alpha_x, \alpha_y, \alpha_z\) for \(i = 1, 2 \text{ and } 3\), respectively. In order for this set of four objects to fulfill these anti-commutation relations, each \(\alpha\) must be, minimally, four-dimensional. One set of matrices which obey similar anti-commutation relations are the Pauli spin matrices: \(\sigma_{x,y,z}\). These are 2x2 matrices, however, and there are only three of them, so they are not useful in their usual form. If the \(\alpha\’s\) are defined as

\[
\alpha_0 = \begin{pmatrix} \mathbb{1}_{2x2} & 0 \\ 0 & -\mathbb{1}_{2x2} \end{pmatrix}
\tag{1.8}
\]

\[
\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}
\tag{1.9}
\]

Where \(\mathbb{1}_{2x2}\) is a two by two identity matrix and \(\sigma_i\) are the Pauli spin matrices:

\[
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\]

Though these four by four matrices do not represent the only set of matrices which satisfy the anti-commutation relations, they may only differ by a similarity transform from this set. Because the Dirac equation contains operators represented by four dimensional matrices, the solutions \(\{\psi\}\) must be represented by a four component vector

\[
\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}
\tag{1.10}
\]

The interpretation of the components of this four-vector is not readily evident. Expressing the Dirac
equation in full matrix form provides some elucidation

\[
\begin{pmatrix}
(p_0 - mc) & 0 & -\hat{p}_z & -\hat{p}_x - i\hat{p}_y \\
0 & (p_0 - mc) & -\hat{p}_x - i\hat{p}_y & \hat{p}_z \\
-\hat{p}_z & -\hat{p}_x - i\hat{p}_y & (p_0 + mc) & 0 \\
-(p_0 + mc) & \hat{p}_z & 0 & (p_0 + mc)
\end{pmatrix}
\begin{pmatrix}
\psi_1 \\
\psi_2 \\
\psi_3 \\
\psi_4
\end{pmatrix} = 0 \quad (1.11)
\]

The non-relativistic electronic wave-function has two components corresponding to the \( \alpha \) and \( \beta \) components of spin angular momentum. In the non-relativistic limit, \( p_0 \) approaches \( mc \), and the terms which couple \( \psi_1 \) with \( \psi_4 \) drop out. What remains are four eigenvector equations, with approximate eigenvalues of \( +m_0c^2 \) for \( \psi_1 \) and \( \psi_2 \), and \( -m_0c^2 \) for \( \psi_3 \) and \( \psi_4 \). \( \psi_1 \) and \( \psi_2 \), then, may be interpreted as the \( \alpha \) and \( \beta \) components of positive energy, electron-like solutions, but the solutions which are dominated by \( \psi_3 \) and \( \psi_4 \) do not possess a readily evident interpretation. Dirac deduced that these solutions correspond to a particle with the same mass as the electron but an opposite charge, and dubbed these particles positrons. Far from being figments of a theorist’s imagination, only three years after Dirac First suggested them in 1930, positrons were observed experimentally.

1.3 Hydrogenic Solutions of Dirac’s Equation

In its present form, the Dirac equation cannot even address the hydrogen atom, since it has no provision for an external potential. If we introduce a scalar and vector potential, \( e\phi + \vec{A} \), to the momenta of the field free Dirac equation

\[
p_0 \rightarrow p_0 - \frac{e\phi}{c} \quad (1.12)
\]

\[
\hat{p} \rightarrow \hat{p} - \frac{\vec{A}}{c} = \vec{\pi} 
\]

then the Dirac equation becomes

\[
\left[p_0 - \frac{e\phi}{c} - \alpha \cdot \vec{\pi} - \beta mc\right] \psi = 0 \quad (1.14)
\]

If the potential terms are not explicitly time dependent, we may make the rearrangement

\[
\left[\beta mc + c\alpha \cdot \vec{\pi} + e\phi\right] \psi = E\psi \quad (1.15)
\]

The left hand side defines the Dirac Hamiltonian for a single particle in a time-independent electromagnetic field. In the absence of a magnetic field the Dirac Hamiltonian, \( \hat{H}_p \), may be expressed in full matrix form as

\[
\hat{H}_p = \begin{pmatrix}
(e\phi + mc^2) & 0 & c\hat{p}_z & c\left(\hat{p}_x - i\hat{p}_y\right)
\\
0 & (e\phi + mc^2) & c\left(\hat{p}_x + i\hat{p}_y\right) & -c\hat{p}_z \\
c\hat{p}_z & c\left(\hat{p}_x + i\hat{p}_y\right) & (e\phi - mc^2) & 0 \\
c\left(\hat{p}_x + i\hat{p}_y\right) & -c\hat{p}_z & 0 & (e\phi - mc^2)
\end{pmatrix} \quad (1.16)
\]

This Hamiltonian is now ready to take on the task of an electron in the field of a nucleus. Because of the spherical symmetry of the nuclear potential, the Schrödinger equation for the hydrogenic atom may be simplified via separation of variables into a radial equation and two angular equations. The angular solutions are given by the spherical harmonics, \( Y_{lm}(\theta, \phi) \), and the
where $L_{n+1}^{2l}(x)$ are the associated Laguerre polynomials, and $n$, $l$, and $m_l$ are the principal quantum number, orbital angular momentum quantum number, and the $z$-projection of $l$, respectively. The hydrogenic solutions to the Dirac equations may also be achieved analytically, though the resulting eigenfunctions cannot be so simply expressed. Because of the presence of the $\vec{\alpha} \cdot \vec{\alpha}$ term, the $\vec{s}$ and $\vec{l}$ operators do not commute with $\tilde{H}$. The components of the $\vec{j} = \vec{l} + \vec{s}$ operator, however, does commute with $\tilde{H}$, and so the eigenfunctions of the Hamiltonian may be eigenfunctions of $\vec{j}^2$ and $\vec{l}^2$ and $\vec{s}^2$ also commute with the Hamiltonian, and so we may still associate a particular $l$ and $s$ value with each solution, $\psi$.

It is useful, at this point, to express the four-component wave-function as a pair of two component spinors, $\phi$ and $\theta$:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \phi \\ \theta \end{pmatrix} \quad (1.18)$$

Such a separation is dictated by the nature of the $4 \times 4$ Pauli spin matrices, and makes it possible to work with the familiar two-spinors of non-relativistic theory. Capitalizing on the commutation of $\tilde{H}$ with $\vec{j}$, each spinor may be expressed as the product of a two-component eigenfunction of $\vec{j}$ which is dependent upon the spin and spatial angular coordinates and a radial function:

$$\phi = g(r) \chi^{i,m_j}_l ; \quad \theta = f(r) \chi^{i,m_j}_l \quad (1.19)$$

The eigenfunctions of $\vec{l}$, $\chi^{i,m_j}_l$, may be expressed as sums of direct products of the more familiar eigenfunctions of $\vec{l}$ and $\vec{s}$ and the appropriate Clebsh-Gordan coefficients:

$$\chi^{j=(1+y^2)m_j}_l = \chi^{(+)}_{j,m_j} = \sqrt{\frac{1 + m_j + y^2}{2l + 1}} Y_{l,m_j+y^2}(\phi y^2) + \sqrt{\frac{1 - m_j + y^2}{2l + 1}} Y_{l,m_j+y^2}(\phi y^2) \quad (1.20)$$

$$\chi^{j=(1-y^2)m_j}_l = \chi^{(-)}_{j,m_j} = -\sqrt{\frac{1 - m_j + y^2}{2l + 1}} Y_{l,m_j-y^2}(\phi y^2) + \sqrt{\frac{1 + m_j + y^2}{2l + 1}} Y_{l,m_j-y^2}(\phi y^2) \quad (1.21)$$

where $Y_{l,m_j=\pm y^2}$ are the spherical harmonics and

$$\phi y^2 = \alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix} ; \quad \phi y^2 = \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.22)$$

Next, it is useful to define the operator

$$\tilde{K} = (\sigma \cdot \vec{l} + 1) \quad (1.23)$$
which commutes with $\tilde{J}$. The eigenvalue equation associated with $\tilde{K}$ is given by

$$
\tilde{K}\chi_{j,m}^{(\pm)} = \pm k \chi_{j,m}^{(\pm)}
$$

(1.24)

where

$$
k = \begin{cases} 
  j + 1/2 & \text{if} \quad j = l - 1/2 \\
  -(j + 1/2) & \text{if} \quad j = l + 1/2
\end{cases}
$$

(1.25)

Because of the spherical symmetry of the nuclear potential, the eigensolutions may be further separated on the basis of their response to spatial coordinate inversion. The parity operator commutes with the full Hamiltonian, and so the final eigenfunctions, $\psi_{j,m}$, must also obey the eigenvalue equation

$$
\beta \psi(-r) = \pm \psi(r)
$$

(1.26)

This necessarily dictates one of the following forms for the four-component eigensolutions:

$$
\psi^{(+)}_{j,m} = \begin{bmatrix} g(r) \chi^{(+)} \\ f(r) \chi^{(-)} \end{bmatrix}
$$

(1.27)

or

$$
\psi^{(-)}_{j,m} = \begin{bmatrix} -g(r) \chi^{(-)} \\ f(r) \chi^{(+)} \end{bmatrix}
$$

(1.28)

It is possible, through the use of the identity

$$
\sigma \cdot \hat{p} = (\sigma \cdot \hat{r}) \left( \hat{r} \cdot \hat{p} + \frac{\sigma \cdot \hat{l}}{r} \right)
$$

(1.29)

and the property

$$
(\sigma \cdot \hat{r}) \psi^{(\pm)}_{j,m} = -\psi^{(\mp)}_{j,m}
$$

(1.30)

to completely separate out the angular functions from the eigenvalue equation, and thereby achieve two coupled, complex differential equations for the radial functions $f(r)$ and $g(r)$. The resultant equations are greatly simplified by the substitutions

$$
f(r) = F(r)/r \quad \text{and} \quad g(r) = iG(r)/r
$$

(1.31)

to yield

$$
\left(mc^2 - E - Z/r\right)F(r) - c \frac{dG(r)}{dr} - kcG(r)/r = 0
$$

(1.32)

$$
\left(-mc^2 - E - Z/r\right)F(r) + c \frac{dG(r)}{dr} - kcG(r)/r = 0
$$

(1.33)

The solutions to these differential equations may be determined by first solving the asymptotic equations in the limit as $r \to \infty$. These solutions, which are of the form
are then multiplied by an undetermined power series for both components and the bound state solutions are sought. The final form of each wave-function is uniquely determined by the recursive relations of the power series and the boundary conditions, but may not be expressed in a simple general form. The resulting electronic energies are given by

\[ E_n = mc^2 \left[ 1 + \frac{(Z\alpha)^2}{\left( n - j - \frac{1}{2} + \sqrt{(j + \frac{1}{2})^2 + (Z\alpha)^2} \right)^2} \right]^{1/2} \]  

(1.36)

where \( \alpha \) is the fine structure constant, and is given by \( 1/c \) in atomic units.

The four-component nature of the Dirac eigenfunctions gives rise to many interesting differences when compared to the non-relativistic solutions. In order to better understand these differences, it is useful to inspect some simple approximations of the lowest energy hydrogenic solutions, for example the \( n = 2, j = \{1/2, 3/2\} \) solutions.

\[
\begin{pmatrix}
(n = 2, j = 1/2, k = 1, m_j = 1/2) \\
(n = 2, j = 1/2, k = 1, m_j = -1/2)
\end{pmatrix}
\begin{pmatrix}
e^{-\rho/2}(1-\rho/2) \\
0 \\
e^{-\rho/2}(1-\rho/2) \\
iZae^{-\rho/2}(1-\rho/4)\cos\theta \\
iZae^{-\rho/2}(1-\rho/4)\sin\theta e^{i\phi} \\
iZae^{-\rho/2}(1-\rho/4)\sin\theta e^{i\phi} \\
0 \\
e^{-\rho/2}(1-\rho/2) \\
-iZae^{-\rho/2}(1-\rho/4)\cos\theta \\
-iZae^{-\rho/2}(1-\rho/4)\cos\theta
\end{pmatrix}

(1.37)
\( n = 2, j = \frac{1}{2}, k = -1, m_j = \frac{1}{2} \)

\[
\frac{1}{4\sqrt{6}} \begin{pmatrix}
e^{-\rho/2} \rho \cos \theta \\
e^{-\rho/2} \rho \sin \theta e^{i\phi} \\
-\frac{3i}{2} Z \rho e^{-\rho/2} (1 - \rho/6) \\
0
\end{pmatrix}
\] (1.39)

\( n = 2, j = \frac{1}{2}, k = -1, m_j = -\frac{1}{2} \)

\[
\frac{1}{4\sqrt{6}} \begin{pmatrix}
e^{-\rho/2} \rho \cos \theta \\
e^{-\rho/2} \rho \sin \theta e^{i\phi} \\
0 \\
-\frac{3i}{2} Z \rho e^{-\rho/2} (1 - \rho/6)
\end{pmatrix}
\] (1.40)

\( n = 2, j = \frac{3}{2}, k = 2, m_j = \frac{3}{2} \)

\[
\frac{1}{8} \begin{pmatrix}
e^{\rho/2} \sin \theta e^{i\phi} \\
e^{-\rho/2} \sin \theta e^{i\phi} \\
-\frac{1}{4} Z \rho e^{-\rho/2} (1 - \rho/4) \sin \theta \cos \theta e^{i\phi} \\
\frac{i}{4} Z \rho e^{-\rho/2} (1 - \rho/4) \sin^2 \theta e^{i\phi}
\end{pmatrix}
\] (1.41)

\( n = 2, j = \frac{3}{2}, k = 2, m_j = -\frac{3}{2} \)

\[
\frac{1}{8} \begin{pmatrix}
e^{\rho/2} \sin \theta e^{i\phi} \\
e^{-\rho/2} \sin \theta e^{i\phi} \\
-\frac{i}{4} Z \rho e^{-\rho/2} (1 - \rho/4) \sin^2 \theta e^{2i\phi} \\
\frac{i}{4} Z \rho e^{-\rho/2} (1 - \rho/4) \sin \theta \cos \theta e^{i\phi}
\end{pmatrix}
\] (1.42)

\( n = 2, j = \frac{3}{2}, k = 2, m_j = \frac{1}{2} \)

\[
\frac{1}{4\sqrt{3}} \begin{pmatrix}
e^{\rho/2} \sin \theta e^{i\phi} \\
-\frac{1}{2} e^{-\rho/2} \sin \theta e^{i\phi} \\
-\frac{i}{8} Z \rho e^{-\rho/2} \rho (1 - 3 \cos^2 \theta) \\
\frac{3i}{8} Z \rho e^{-\rho/2} \rho \sin \theta \cos \theta e^{i\phi}
\end{pmatrix}
\] (1.43)
\[
\left( n = 2, j = 3/2, k = 2, m_j = -1/2 \right)
\]
\[
\frac{1}{4\sqrt{3}} \left( \frac{1}{2} e^{-\rho^2/2} \sin \theta e^{i\phi} + \frac{1}{2} e^{-\rho^2/2} \sin \theta e^{i\phi} \right)
\]
\[
\frac{3i}{8} Z e^{-\rho^2/2} \rho \sin \theta \cos \theta e^{i\phi}
\]
\[
\frac{i}{8} Z e^{-\rho^2/2} \rho \left( 1 - 3\cos^2 \theta \right)
\]

These solutions are correct to order \( Z\alpha \), as illustrated in Moss's discussion of the hydrogenic solutions to Dirac's equation [R. E. Moss, Advanced Molecular Quantum Mechanics (John Wiley and Sons, New York, 1973)]. Here, analogous to the non-relativistic case, for each unique combination of the principal quantum number \( n \) and a specific angular momentum quantum number \( j \), the solutions must be further classified according to the projection of their total angular momentum on the \( z \)-axis, \( m_j \). These solutions are, of course, degenerate in the absence of an external field. The contributions from the first two components of the wave-function tend to be much larger than the contributions from the final two components for electron-like solutions. For this reason the upper two-component spinor is known as the large component of the wave-function, while the lower spinor is known as the small component. By orthogonality, the opposite is true for the positron-like solutions. Since the angular and spin portions of the wave-functions are normalized, the magnitude of the large and small contributions is determined entirely by the radial functions, \( f(r) \) and \( g(r) \). Given the asymptotic form of the radial equations (1.34) and the orbital energy (1.36), the ratio of their contributions may be expressed as

\[
g(r) \approx \frac{Z\alpha}{2n} \quad f(r) \quad (1.45)
\]

In general, the ratio will probably be larger than this approximate value in regions close to the nucleus, where the contribution of \( g(r) \) is the greatest. The contribution of the small component, then, may be significant for sufficiently heavy nuclei. Though hydrogenic ions with very heavy nuclei may not be of great practical interest, this relationship will have implications for heavy many-electron atoms where the innermost electrons experience a large portion of the full nuclear charge and hence are closely related to the analogous hydrogenic systems.

One interesting consequence of the four-component wave-function may be observed in the associated probability density. Scalar wave-functions, \( \psi \), which possess radial and angular nodes will have the same radial and angular nodes in their probability density, \( \psi^* \psi \). The four-component wave-function, however, gives rise to a probability density which is the sum of the probability densities associated with the large and small wave-functions, \( \phi \) and \( \chi \), respectively. The radial component of \( \phi \) and \( \chi \) possess different numbers of nodes and, in general, none of these nodes will coincide. Similarly, the angular contributions of \( \phi \) differ from those of \( \chi \) by a single unit of angular momentum. Therefore, although \( \phi^* \phi \) and \( \chi^* \chi \) may possess either radial of angular nodes individually, their sum, their sum will be node-less.

The energy eigenvalues of the hydrogenic solutions to the Schrödinger equation are only dependent upon \( n \), the principal quantum number, while the Dirac hydrogenic eigenvalues are dependent on both \( n \) and \( j \). The spectral dependence on \( j \) is upheld by experimental observation, however the degeneracy of eigenvalues for solutions with the same \( j \) values but differing \( l \) values is not observed in nature. The breaking of this degeneracy is known as the Lamb shift and its origin has been attributed to the difference in the interaction of the different \( j \) eigenfunctions with the vacuum fluctuations predicted by quantum electrodynamics. The magnitude of the Lamb shift is much
smaller than the splitting introduced by spin orbit coupling and is largest for the \( n = 2, j = 1/2 \) shells of the hydrogenic atom. In this case, the splitting introduced by the Lamb shift is approximately 10\% of the magnitude of the energy separation of the \( p_{1/2} \) and \( p_{3/2} \) eigenfunctions. It should not be too surprising that the predictions of the Dirac equation cannot entirely reproduce experimental observation, since it has its roots in special relativity. In order for any physical theory to definitively represent reality, it must at least be consistent with the general theory of relativity.

1.4 Electron Spin Angular Momentum and Magnetic Moment

One of the triumphs of the Dirac equation was its explicit connection to electron spin. The Dirac equations does not predict electron spin as a relativistic property, observable only under relativistic conditions. Reduction of the Dirac equation for an electron in a magnetic field to its non-relativistic limit yields the Schrödinger equation with a correction term which takes account of the interaction of the electron’s intrinsic magnetic moment with the magnetic field. Using the definition of the Dirac Hamiltonian, \( \hat{H}_D \), given in (1.15), we may arrive at the equation

\[
\left( \hat{H}_D + e\phi \right)^2 / c^2 = (\beta mc + \alpha \cdot \pi)^2
\]

(1.46)

Due to the anti-commutation properties of the components of \( \alpha \), we may expand the right hand side to reveal that

\[
\left( \hat{H}_D + e\phi \right)^2 c^{-2} = m^2 c^2 (\alpha \cdot \pi)(\alpha \cdot \pi)
\]

(1.47)

If we define our \( \alpha_i \) matrices as products of a matrix \( \rho \) and the \( \sigma_{ia} \) matrices where

\[
\rho = \begin{pmatrix} 0 & 1_{2 \times 2} \\ 1_{2 \times 2} & 0 \end{pmatrix} \quad \text{and} \quad \sigma_{i4} = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}
\]

(1.48)

so that \( \alpha_i = \rho \otimes \sigma_{ia} \), then we may express \((\alpha \cdot \pi)(\alpha \cdot \pi)\) as

\[
(\alpha \cdot \pi)(\alpha \cdot \pi) = (\sigma \cdot \pi)(\sigma \cdot \pi)
\]

(1.49)

where we have used the identity \( \rho^2 = 1 \). This may be rearranged to the form

\[
\begin{align*}
\rho 
\end{align*}
\]

Radially weighted \( f(\tau), g(\tau) \) and \( \rho(\tau) \) for the 2s\( \frac{3}{2} \) orbital of Xe
\[(\sigma \cdot \pi)(\sigma \cdot \pi) = \pi \cdot \pi + i\sigma(\pi \times \pi) \quad (1.50)\]

via sundry vector identities. The cross-product of \( \pi \) with itself does not drop out. This is because \( \pi \) is the sum of two vectors, \( p \) and \( A \), so the cross terms of the cross products remaining, making the cross-product of \( \pi \) with itself

\[(\pi \times \pi) = -i\hbar \nabla \times A = -i\hbar B \quad (1.51)\]

Returning to our original equation (1.39), we may make this substitution to arrive at

\[
\left(\hat{H}_D + e\phi\right)^2 c^{-2} = m^2 c^2 + \pi^2 + h\sigma_i B
\]

(1.52)

In the non-relativistic limit this gives us an Hamiltonian of the form

\[
\hat{H}_D = -e\phi + \frac{\pi^2}{2m} + \frac{\hbar}{2m} \sigma_i B
\]

(1.53)

The final term indicates that a term which corresponds to a body with a magnetic moment of \(-\frac{\hbar}{2m}\sigma\) in a magnetic field \(B\) must be added to give the proper energy. This prediction of electron spin as a property which should exist in the non-relativistic limit provides strong evidence that Dirac’s equation provides a more complete physical picture than the wave equation of Schrödinger, where electron spin must be treated in an \textit{ad hoc} manner.