Dissertation for the degree
of Doctor Scientiarium
Trond Saue

Principles and Applications of Relativistic Molecular Calculations

Department of Chemistry

Faculty of Mathematics and Natural Sciences

University of Oslo 1996

Contents

List of papers

1. T. Saue, T.Helgaker, K. Fægri and O.Gropen:
   Principles of direct 4-component relativistic SCF:
   Application to dihydrides of group 16
   Manuscript.

2. T. Saue and H.J.Aa. Jensen:
   Quaternion symmetry in relativistic molecular calculations:
   I. The Dirac-Fock method
   Submitted to J. Chem. Phys.

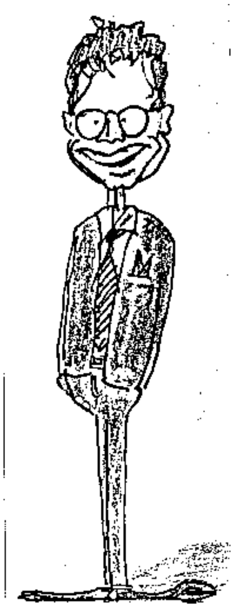
3. H.J.Aa.Jensen, K.G.Dyall, T.Saue and K.Fægri:
   Relativistic 4-component Multi-Configurational Self-Consistent Field Theory for Molecules: Formalism
   Accepted for publication in J.Chem.Phys.

4. L.Visscher, T.Saue, W.C.Nieuwpoort, K.Fægri and O.Gropen:
   The electronic structure of the PtH molecule:
   Fully relativistic configuration interaction calculations
   of the ground and excited states
   J.Chem.Phys. **99** (1993) 6704–6715

5. T.Saue, K.Fægri and O.Gropen:
   Relativistic effects on the bonding of heavy and superheavy hydrogen halides
   Submitted to Phys.Rev.Lett.

Acknowledgments

> *Looking back over my whole life in research, I view the three years after graduation as providing an extremely valuable foundation. The competitive swimmer glides under water for a brief interval after he has dived — it was that kind of preparatory period for me.*
>
> HIDEKI YUKAWA*[1]*



This thesis is based on research carried out during the years 1991 – 1995, to a large extent at the Department of Chemistry at the University of Oslo. I spent the year September 1994 to September 1995 as an assistant lecturer at the Department of Chemistry at the Univeristy in Tromsø rediscovering my north-Norwegian roots. I have furthermore had some very rewarding stays at the University of Odense, Denmark, and at the Univerisity of Groningen, The Netherlands. My thanks go to the universities for their hospitality. Economical support from The Norwegian Research Council (NFR) and the REHE program of the European Science Foundation is gratefully acknowledged.

There are a number of people, whose support and encouragement has proven invaluable to me during these years. First and foremost I would like to thank my supervisors, Knut Fægri jr. and Odd Gropen, for all their support and good will, and for trying to keep me on the straight and narrow path. I would like to single out a few other people who have meant a lot:

- Ken Dyall, Hans J.Aa.Jensen and Luuk Visscher, for an extremely rewarding cooperation

- Trygve Helgaker, for lots of enthusiasm, creativity and fun

- Jon K. Lærdahl, for your thouroughness, help and friendship

- my landlady, Eva J. Moestue, for being a marvelous person
- Kenneth Ruud, for trying to teach my about magnetic properties and the viola. . .
- Merethe Sjøvoll, my comrade in arms. Good luck with your thesis !!!
- Ole Swang, for strong personal support whenever needed
- my family, for lots of love

In addition I have learned a lot from interchange with J.Almlöf, J.Asplin, G.Aucar, E.Engel, E.K.U.Gross, B.A.Heß, J.-L. Heully, U.Kaldor, J.Oddershede, J.Olsen, P.Pyykkö, H.Quiney, L.Seijo, C.Teichteil, L.Veseth and U.Wahlgren.

Thanks also to the students (past and present) of the groups of theoretical chemistry at the Universities of Oslo and Tromsø for lots of fun and memories: Kenneth Ruud, Hege Strømsnes, Vebjørn Bakken, Hanne Heiberg, Kjerstin Ellingsen, Torgil Vangeberg, Hilde Fagerli, Geir Edvardsen, Ståle Gove, Hege Ryeng and Elly Karlsen. Let us keep in touch !

I have also had the pleasure of frequenting the famous Café Erwin . . .

Oslo, December 1995

Trond Saue

Introduction

At his first lecture on quantum mechanics in Michaelmas term 1945, Dirac entered the room to find it jammed with students. He obviously did not expect to see so many students at his lecture because he announced 'This is a lecture on quantum mechanics.' No one moved. He repeated the announcement. Nothing happened again.

S.SHANMUGADHASAN *(1987) [2]*

0.1 General overview

Even though P. A. M. Dirac, the originator of the relativistic wave equation for the electron, stated that relativistic effects were[3]

> of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions,

it has become increasingly clear during the past few decades that in many areas of chemistry such effects may not be neglected [4, 5]. Discussions of relativistic effects are now entering basic chemistry textbooks [6], and there is a rapidly expanding literature on the subject [7, 8, ?]. This acknowledgement stems largely from the extensive progress in the development of methods and technology which has made chemical systems containing heavy element atoms accessible for computational studies. It is found that the non-relativistic Schrödinger-equation fails to give an adequate description of such systems. A well known example is the relativistic effects on the band structure of metallic gold[9, 5]. Non-relativistic calculations overestimates the gap between the 5d and 6s band and predicts absorption in the UV region, which would give gold an appearance similiar to that of silver. In atoms, relativity generally leads to a contraction of s and p orbitals and (indirectly) to an expansion of d and f orbitals. In addition the spin-orbit coupling causes the fine structure of atomic spectra. The effect of relativity in molecular systems is more uncertain and is currently an area of active research. The interest in the effects of relativity sparked in 1992 the establishment of the programme "Relativistic Effects in Heavy Element Chemistry and Physics"(REHE), sponsored by the European Science

Foundation [10]. The goal of this programme is to spur interest and facilitate cooperation among European scientists in the field.

The starting point of relativistic molecular calculations is the one-electron Dirac operator in the external field of fixed nuclei [11, 12, 13]. The two-electron interaction is usually represented by the standard Coulomb interaction. This operator is clearly not Lorentz invariant, but may be considered as the zero-order term in a series expansion of the full two-electron interaction, which is not available in closed form. First order corrections may be added in the form of the Gaunt or Breit terms. The effect of these corrections on molecular properties is not very well understood, but they appear to be small. The resulting Dirac-Coulomb(-Gaunt/Breit) Hamiltonian leads to a considerable increase in the computational effort as compared to the corresponding non-relativistic electronic Hamiltonian. This is due to the fact that that the Dirac operator contains spin and leads to a 4-component rather than scalar wave function. In addition the wave function can generally not be chosen real, thus forcing the use of complex algebra. Consequently, several approximative approaches have been investigated. These approaches are usually based on a transformation of the Dirac operator to a truncated two-component form and a subsequent separation of spin-free and spin-dependent terms [14]. The spin-free terms (e.g. the spin free no pair Douglas-Kroll operator [15]) are straightforwardly incorporated in conventional methods and codes, whereas spin-dependent terms may be added at a latter stage in the calculations, as in spin-orbit CI[16] Another possibility is to use the Dirac-Coulomb Hamiltonian or quasi-relativistic operators to generate relativistic effective core potentials(RECP) [17, 18, 19, 20].

During the past 15 years several 4-component molecular Dirac-Fock codes have been developed[21, 22, 13, 23, 24, 25]. The first basis set calculations were flawed due to the disregard of the coupling between the large and small components which must be reflected in any basis set expansion [26, 27]. Also, there has been considerable confusion over the variational foundation of the method[11, 28]. The Dirac-Coulomb Hamiltonian generates both positron and electron solutions, so that the electronic ground state is an excited state in its spectrum. Therefore the minimalization of the corresponding non-relativistic method must be replaced by a minimax principle. Convergence is straightforwardly obtained by vector selection, but may be more rigorously procured by second order methods.

The one-particle basis generated by the Dirac-Fock method may be employed in correlated methods to obtain more accurate results. Implementations of the second order Møller-Plesset(MP2) [29, 30], the multi-reference Configuration Interaction (CI) [22, 31] and coupled-cluster singles and doubles (CCSD) [32] methods have been reported. Also, work is in progress on the development of a 4-component molecular Multi-Configurational Self-Consistent Field (MCSCF) code[33], as presented in this thesis. Relativity has furthermore been approached by means of density functional theory [34]. Since relativistic effects are predominantly found in systems containing heavy elements with a large number of electrons in the valence region, the effects of dynamic correlation may be pronounced. Static correlation reflects the need for a multiconfigurational wave function in near-degeneracy situations, such as bond-breaking. In relativistic systems such a situation becomes particularly critical due to the additional near degeneracies provided by the spin-orbit coupling [16]. In fact, the breaking of spin symmetry by the spin-orbit coupling generally makes a multiconfigurational approach mandatory for systems with more than one open shell [33].

The motivation for the development of molecular 4-component methods is manifold. The study of relativistic effects has already been mentioned. The prime motivation, however, is to obtain computational tools that allow an adequate description of chemical systems in which such effects are important. In addition, 4-component calculations serve as benchmarks for the quasirelativistic methods and probe their range of validity. In addition, the 4-component methods are particularly well suited for studies of electric and magnetic properties of molecules, due to the simple structure of operators. As an example one may consider nuclear spin-spin coupling, where four operators in the non-relativistic formalism (Fermi contact, spin-dipole, paramagnetic spin-orbit and diamagnetic spin-orbit) are replaced by one operator in the 4-component formalism [35]. At present, the computational intensity of the 4-component methods to some extent limits feasible applications. On the other hand, this has lead to intensive work on the computational methods, work which in the end may benefit non-relativistic methods as well.

0.2 Layout of the thesis

The thesis presented here focuses on the methodological aspects of relativistic molecular calculations. In particular, it presents the formalism for

- the quaternion Dirac-Fock equations
- the direct 4-component Dirac-Fock method
- the multi-configurational self-consistent field (MCSCF) method

Applications are represented by

- 4-component configuration interaction (CI) studies of the five lower states of PtH
- 4-component direct SCF studies of bonding in hydrides of iodine, astatine and eka-astatine (element 117)
- 4-component direct SCF studies of bonding in dihydrides of tellurium, polonium and eka-polonium (element 116)

Methodological development depends on a clear understanding of both the mathematical structure and the physical content of the theory to be implemented. *This has been a decisive factor in the layout of this thesis.*
The thesis consists of three parts.

1. In the first part I give an extensive, but not comprehensive, introduction to the field of relativistic quantum mechanics. I have tried to bring out the mathematical structure and physical implications of the Dirac equation and its approximative many-electron extension, the Dirac-Coulomb Hamiltonian. The presentation is interspersed with some of my own research material, where I have felt that it elucidates the presentation. I have also tried to avoid excessive overlap with the papers of the second part, which means that the two parts should be read as a whole.
2. In the second part five papers are presented.
3. I have furthermore included several appendices:
    - documentation of the 4-component direct SCF program DIRAC
    - details on the reduction of the Breit term to two-component form
    - background material on symmetry in relativistic systems
    - diagonalization of quaternion Hermitian matrices
    - tabulation and visualization of the angular part of atomic 2-spinors

Notation: I use atomic units throughout, but write electron mass m and the speed of light c out explicitly.

Part I

Relativistic Quantum Mechanics

Chapter 1

One-electron systems

There was a young lady named Bright
Whose speed was far faster than light
She set off one day
In a relative way
And returned home the previous night

A. H. R. BULLER*[36]*

The main objective of this chapter is to obtain a Lorentz invariant description of an electron in a molecular field, that is the field of nuclei. We outline central features of the theory of special relativity and consider the transition from classical to quantum mechanics. We present the Dirac equation and discuss its mathematical structure and physical content. The two quantities are inextricably connected:

1. The transition from classical to quantum mechanics usually proceeds by way of analogy. Spin has no classical analogue, but appears explicitly in the Dirac equation. The spin is coupled to the spatial degrees of freedom, and this has profound consequences for the symmetry properties of solutions to the Dirac equation. The solutions are fermion functions and change sign under a rotation 2π. This feature is usually accounted for by the introduction of double groups. The behaviour of inversion in double group theory is deduced on the basis of classical analogies. We shall demonstrate that by deriving an explicit representation of inversion in spin space, we arrive at a contradictory result.

2. The coupling of spin and spatial degrees of freedom means that the spin symmetry of non-relativistic theory is lost. In the absence of external magnetic fields, however, the spin symmetry can to some extent be replaced by time reversal symmetry. We demonstrate that time reversal symmetry yields a quaternion formulation of the Dirac equation.

3. The charge conjugation symmetry of the free particle Dirac equation reveals its many-body aspect in that it describes both the electron and its antiparticle, the positron. A proper decription of an electron in an external field can only be obtained within the framework of quantum electrodynamics (QED).

Further information of the physical contents of the Dirac equation is obtained by considering perturbation expansions in terms of the fine structure constant α. This is also the way to quasirelativistic one- and two-component Hamiltonians. We discuss the difficulties involved in their derivation, namely the risk of introducing unbounded and/or highly singular operators.

1.1 Special relativity

> *You are rowing a boat upstream. The river flows at three miles per hour; your speed against the currents is four and one-quarter. You lose your hat on the water. Forty-five minutes later you realize it is missing and execute the instantaneous, acceleration-free about-face that such puzzles depend on. How long does it take to row back to your flowing hat?*[1]
>
> JAMES GLEICK *(1992) [37]*

The theory of special relativity is "special" because it only considers the transformation of space- and time coordinates between inertial frames, that is coordinate systems in uniform relative motion. The *principle of relativity*, however, predates the theory set forward in 1905 [38] by Albert Einstein, technical expert third class at the patent office in Bern, and states that [39]

> The laws of physics take the same form in all inertial frames.

Implicit in this postulate is the assumption of homogeneity of space and time and isotropy of space. Considerations of the structure of time and space form a powerful tool for the elucidation of the laws of physics and their mathematical formulation , as is demonstrated in section 1.2 and 1.4. The principle of relativity was originally connected to the notion of absolute time as embodied in the *Galilean transformation*

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} - \mathbf{v}t \\ t' &= t \end{aligned} \tag{1.1}$$

relating coordinates of inertial frames K$'$ and K, having the same orientation of axes, and where K$'$ moves with uniform velocity $\mathbf{v}$ relative to K. In the theory of special relativity the idea of absolute time is replaced by the postulate [39]:

> In any given inertial frame, the velocity of light c is the same whether the light be emitted by a body at rest or by a body in uniform motion.

This postulate leads directly to the *Lorentz transformation*, as follows: The speed of light, as measured in inertial frames K and K$'$, shall have the same value

$$\frac{|\mathbf{r}_2' - \mathbf{r}_1'|}{t_2' - t_1'} = c; \qquad \frac{|\mathbf{r}_2 - \mathbf{r}_1|}{t_2 - t_1} = c \tag{1.2}$$

---

[1] "A simpler problem than most. Given a few minutes, the algebra is routine. But a student whose head starts filling with 3s and $4\frac{1}{4}$s, adding them or subtracting them, has already lost. This is a problem about reference frames. The river's motion is irrelevant — as irrelevant as the earth's motion through the solar system or the solar system's motion through the galaxy. In fact all the velocities are just so much foliage. Ignore them, place your point of reference at the floating hat — think of yourself floating like the hat, the water motionless about you, the banks an irrelevant blur — now watch the boat, and you see at once [. . .] that it will return in the same forty-five minutes it spent rowing away." [37]

By simple rearrangement we find that the quadratic form

$$s_{12}^2 = (\mathbf{r}_2 - \mathbf{r}_1)^2 - c^2t^2 = \left(\mathbf{r}_2' - \mathbf{r}_1'\right)^2 - c^2t'^2 \tag{1.3}$$

is identically zero in both coordinate systems. From the assumption of homogeneity of time and space and isotropy of space, one can deduce [40] that the *interval* s_{12} between *any* two sets of space and time coordinates (events) is conserved under the transformation between inertial frames. To simplify things a bit, we shall let frames K and K$'$ coincide at time $t_1 = t_1' = 0$ so that $\mathbf{r}_1' = \mathbf{r}_1 = 0$. The interval Eq.(1.3) may then be thought of as the length of a four-dimensional vector r_μ which shall be denoted *4-position*

$$r_\mu = (r_1, r_2, r_3, r_4) = (\mathbf{r}, ict) \tag{1.4}$$

The Lorentz transformation preserves the length of 4-position and is a rotation in 4-dimensional *space-time* where the timelike coordinate is given by $r_4 = ict$. [2] To derive the explicit form of the Lorentz transformation we first consider the transformation between K$'$ and K where $\mathbf{v}$ is directed along the z-axis of both systems, so that x- and y-coordinates can be neglected. Since length of 4-position is conserved the coordinates (z', ict') of K$'$ and (z, ict) of K are related by an orthogonal transformation, which has the general form

$$\begin{bmatrix} z' \\ ict' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} z \\ ict \end{bmatrix} \tag{1.5}$$

Alternatively, we can write out the two equations

$$\begin{aligned} z' &= z\cos\theta + ict\sin\theta \qquad (a) \\ ict' &= -z\sin\theta + ict\cos\theta \qquad (b) \end{aligned} \tag{1.6}$$

Since K$'$ and K coincided at $t_1 = t_1' = 0$ the origin of K$'$, $z' = 0$, has coordinate $z = vt$ at time t in K. Insertion in Eq.(1.6a) immediately gives

$$\tan\theta = \frac{iv}{c} \quad \Rightarrow \quad \cos\theta = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = \gamma, \quad \sin\theta = \frac{iv\gamma}{c} \tag{1.7}$$

from which we obtain the transformation

$$\begin{bmatrix} z' \\ ict' \end{bmatrix} = \begin{bmatrix} \gamma & \frac{iv\gamma}{c} \\ \frac{-iv\gamma}{c} & \gamma \end{bmatrix} \begin{bmatrix} z \\ ict \end{bmatrix} \tag{1.8}$$

---

[2] The Lorentz transformation is only a rotation when translations are excluded from the space-time transformations. With translations included the Lorentz transformation is denoted inhomogeneous and only the distance between 4-position vectors, as expressed by the interval, is conserved in the transformation between inertial frames.

In order to bring the transformation on the same form as Eq.(1.1) we first write

$$\mathbf{r}' = \mathbf{r} + [(\gamma - 1)\, z - \gamma v t]\, \frac{\mathbf{v}}{v}; \qquad t' = \gamma t - \frac{\gamma v z}{c^2} \tag{1.9}$$

and then use the relation $\mathbf{v} \cdot \mathbf{r} = vz$ [41] to obtain

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} + \mathbf{v}\left[(\gamma - 1)\,\tfrac{\mathbf{r}\cdot\mathbf{v}}{v^2} - \gamma t\right] \\ t' &= \gamma\left[t - \tfrac{\mathbf{v}\cdot\mathbf{r}}{c^2}\right] \end{aligned} \tag{1.10}$$

Note that in the limit of $c \rightarrow \infty$ (the non-relativistic limit) we recover the Galilean transformation. The Lorentz (and Galilei) transformation was derived with the restriction that the orientation of spatial axes were identical in K and K′[3], which means that the transformation Eq.(1.10) does not describe all possible rotations in space-time. For instance, a rotation about the time-like axis corresponds to a rotation in ordinary three dimensional space.

The set of 4×4 rotation matrices in space-time forms a continuous group, the (homogeneous) Lorentz group. Four-dimensional vectors, such as 4-position, whose length is preserved under Lorentz transformations, are denoted *4-vectors*. Another 4-vector that we shall make use of is the *4-gradient*

$$\partial_\mu = \left(\boldsymbol{\nabla}, -\frac{i}{c}\frac{\partial}{\partial t}\right) \tag{1.11}$$

The 4-vectors form a convenient and compact formalism for the construction of Lorentz invariant mechanics in analogy with the Newtonian (Galilean invariant) mechanics. We shall do so, in a somewhat heuristic manner, where our ultimate goal is to derive an expression for the energy of an electron in an external field (e.g. the field of nuclei), which will then be the starting point for the transition to relativistic quantum mechanics. A basic prerequisite for such a derivation is a Lorentz-invariant time-like quantity. We define *proper time* τ by

$$d\tau = \frac{ds}{c} = dt\left[1 - \frac{\left(\frac{dx^2}{dt^2} + \frac{dy^2}{dt^2} + \frac{dz^2}{dt^2}\right)}{c^2}\right]^{\frac{1}{2}} = dt\left[1 - \frac{v^2}{c^2}\right]^{\frac{1}{2}} = \gamma^{-1} dt \tag{1.12}$$

The coordinates $\mathbf{r}$ may be thought of as the coordinates of a particle moving with velocity $\mathbf{v}$. In the rest frame of the particle we have $v = 0$ so that $d\tau = dt$. Proper time is thus seen to be the time in the rest frame of the particle.

---

[3] This particular kind of Lorentz transformations is denoted a boost.

Differentation of 4-position with respect to proper time gives *4-velocity*

$$v_\mu = \frac{dr_\mu}{d\tau} = \gamma\left(\mathbf{v}, ic\right) \tag{1.13}$$

Note that the presence of γ in the space-part of 4-velocity limits all speeds to values equal to or below the speed of light. This has the important consequence that there can be no instantaneous interactions. Clearly then, the standard Coulomb potential is not Lorentz-invariant and must be replaced by a potential that take retardation of the interaction into account.

We proceed by defining *4-momentum* as

$$p_\mu = mv_\mu = \left(\mathbf{p}, iMc\right); \qquad \mathbf{p} = M\mathbf{v}, \quad M = \gamma m \tag{1.14}$$

The latter relation indicates that the mass M of a particle increases with the speed [4] The timelike component of p_μ can be associated with energy. To demonstrate this, we first define *4-force* by differentation of 4-momentum with respect to proper time

$$F_\mu = \frac{dp_\mu}{d\tau} = \gamma\left(\mathbf{F}, ic\frac{dM}{dt}\right) \tag{1.15}$$

Since p_μ is a 4-vector its length must be Lorentz invariant. We find

$$\left(p_\mu\right)^2 = M^2\left(v^2 - c^2\right) = -(mc)^2 \tag{1.16}$$

Differentation of Eq.(1.16) with respect to proper time gives the important relation

$$\frac{d\left(p_\mu\right)^2}{d\tau} = 2p_\mu\frac{dp_\mu}{d\tau} = 0 \quad \Rightarrow \quad F_\mu p_\mu = 0 \tag{1.17}$$

showing that 4-momentum and 4-force are orthogonal 4-vectors. A simple rearrangement gives

$$\frac{dMc^2}{dt} = \mathbf{F}\cdot\mathbf{u} = \frac{\mathbf{F}\cdot d\mathbf{r}}{dt} = \frac{dE}{dt} \tag{1.18}$$

where we have used the classical definition of work and energy in the final step. This relation shows that the time derivative of the quantity Mc^2 is associated with the time derivative of energy. We perform a bold generalization:

$$d(Mc^2) = dE \quad \Rightarrow \quad E = Mc^2 \tag{1.19}$$

---

[4] Alternative views are expounded in [42, 43].

The above relation shows that an infinitesimal increase in the kinetic energy of a particle gives a proportional increase in the mass of the particle, so that mass and energy are equivalent. We can now write the 4-momentum of a free particle as

$$p_\mu = \left(\mathbf{p}, \frac{iE}{c}\right) \tag{1.20}$$

By insertion into Eq.(1.16) we finally obtain

$$E^2 = m^2c^4 + c^2p^2 \tag{1.21}$$

This is the relativistic expression for the energy of a free particle.

We seek, however, the description of an electron in a molecular field. External electromagnetic fields are introduced by means of the substitution [40]

$$p_\mu \rightarrow \pi_\mu = p_\mu - qA_\mu; \quad A_\mu = \left(\mathbf{A}, \frac{i}{c}\phi\right) \tag{1.22}$$

where we have introduced the *4-potential* A_μ and the charge q of the particle. The vector and scalar potentials **A** and ϕ are related to the electric and magnetic fields **E** and **B** by

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}; \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{1.23}$$

There exists a many-to-one correspondence between electromagnetic potentials and fields, in that the electric and magnetic fields are invariant under the *gauge* transformation

$$A_\mu \rightarrow A_\mu - \partial_\mu f \tag{1.24}$$

where f is any scalar function of space and time coordinates. For consistency we must therefore require that the laws of physics are invariant under gauge transformations. This is ensured by the substiution Eq.(1.22).

The fact that the 4-potential transform as the 4-position has some very important consequences. Consider a intertial frame K in which there is a scalar potential $\phi(\mathbf{r})$, but no vector potential and thus no magnetic field. By transforming to an inertial frame K′ moving with uniform velocity **v** relative to K, we find a nonzero vector potential

$$\mathbf{A}' = -\frac{\gamma}{c^2}\phi(\mathbf{r})\mathbf{v} \tag{1.25}$$

Note that the vector potential in K′ is expressed in terms of a function of coordinates of K. By transformation of the coordinates to those of K′ retardation terms appear in the vector potential. We shall not consider these explicitly. For $v \ll c$, the resulting magnetic field in K′ is approximated by

$$\mathbf{B}' = -\frac{\mathbf{E} \times \mathbf{v}}{c^2} \tag{1.26}$$

This means that a particle moving in an electric field experiences a magnetic field in its own rest frame. This is the origin of the *spin-orbit effect* [44].

If we perform the substitution Eq.(1.22) and take the square root on both sides in the resulting equation, we finally arrive at a Lorentz invariant energy expression for a particle in an external field

$$E = \pm\sqrt{m^2c^4 + c^2\pi^2} + q\phi \tag{1.27}$$

We see that we have a choice of whether to take the positive or negative root. In classical mechanics, the positive root is chosen, and the negative energy solutions discarded since discontinous changes are not allowed. In order to arrive at the non-relativistic limit, we readjust our energy scale by subtracting the rest mass term mc^2 and expand the square root in $(\pi/mc)^2$

$$\begin{aligned} E' = E - mc^2 &= q\phi + mc^2\left[1 + \frac{\pi^2}{2m^2c^2} - \frac{\pi^4}{8m^4c^4} + \dots\right] - mc^2 \\ &= \frac{\pi^2}{2m} + q\phi + O\left(\frac{\pi^4}{m^3c^2}\right) \end{aligned} \tag{1.28}$$

The first two terms constitute the non-relativistic energy expression for a particle in an exteral field. We shall see in section 1.8, however, that the expansion of the sqaure root in Eq.(1.27) in order to obtain approximations to the relativistic energy is beset with difficulties.

1.2 Spin

> *A colleague who met me strolling rather aimlessly in the beautiful streets of Copenhagen [1922] said to me in a friendly manner,"You look very unhappy"; whereupon I answered fiercely,"How can one look happy when he is thinking about the anomalous Zeeman effect?"*
>
> WOLFGANG PAULI *(1945) [45]*

In the previous section we derived the Lorentz invariant energy expression for a particle in an external field in terms of classical mechanics. We now consider the transition from classical to quantum mechanics. The transition usually proceeds by way of analogy. The intrinsic angular momentum of the electron (spin) was, however, introduced with no classical analogue in the early days of quantum mechanics and therefore deserves special consideration.

We first briefly recall the formalism of quantum mechanics. In the Hilbert-space formalism the states of a system are represented by unit vectors Ψ in an infinite dimensional function space, the Hilbert space. A *scalar product* is defined by

$$\langle \Psi_i \mid \Psi_j \rangle = \int \Psi_i(\tau)^\dagger \, \Psi_j(\tau) \, d\tau \tag{1.29}$$

where τ are coordinates and $d\tau$ the associated volume element. The state vectors are normalized to unity

$$\langle \Psi \mid \Psi \rangle = 1 \tag{1.30}$$

Observables of the system are represented by Hermitian operators in the Hilbert space. The function space is *complete* in the sense that any state vector can be expanded in the eigenvectors Φ of any operator $\hat{\Omega}$ corresponding to some observable:

$$\left\{ \Phi_i | \hat{\Omega} \Phi_i = \omega_i \Phi_i \right\} \quad \Rightarrow \quad \Psi = \sum_i c_i \Phi_i; \quad c_i = \langle \Phi | \Psi \rangle \tag{1.31}$$

The square of the expansion coefficients c_i gives the probability of the corresponding value ω_i of the observable. The expectation value of the observable is

$$\left\langle \hat{\Omega} \right\rangle = \left\langle \Psi \left| \hat{\Omega} \right| \Psi \right\rangle = \sum_i \omega_i c_i^2 \tag{1.32}$$

Another quantity related to experiment is the transition probability

$$\left| \left\langle \Psi_j \left| \hat{\Omega} \right| \Psi_i \right\rangle \right|^2 \tag{1.33}$$

which gives the probability of transition from the state Ψ_i to another state Ψ_j under the influence of an interaction represented by $\hat{\Omega}$. The normalization of state vectors leaves an arbitrary unimodular complex phase undetermined. The state vectors are therefore more properly considered as unit *rays* in the Hilbert space. The absolute phase of the wave function can not be observed experimentally since the phases enter neither probabilites nor expectation values. Relative phases can, however, be observed by interferometry (see e.g. [46, 47, 48]) [5].

Operators are traditionally derived by correpondence with classical mechanics, e.g. the free-particle Schrödinger equation is obtained from the corresponding classical expression

$$E = \frac{p^2}{2m} \quad \Rightarrow \quad \hat{H}\Psi = i\frac{\partial \Psi}{\partial t}; \qquad \hat{H} = -\frac{1}{2m}\nabla^2 \tag{1.34}$$

by the operator substitution

$$E = -icp_4 \rightarrow i\frac{\partial}{\partial t}; \qquad \mathbf{p} \rightarrow i\boldsymbol{\nabla} \tag{1.35}$$

A quantity with no classical analogue is the electron spin, whose existence was first postulated in 1925 by Goudsmit and Uhlenbeck [50]. Spin appeared for the first time in a wave equation,without having been inserted by hand, in the equation Dirac obtained by linearizing the operator analogue of the relativistic energy expression for a free particle Eq.(1.21), and was therefore taken to be a relativistic effect. However, as he stated himself, Dirac was "just playing with equations" [51] and seeing what they gave. In particular he tried to explore the relation

$$(\boldsymbol{\sigma} \cdot \mathbf{p})\,(\boldsymbol{\sigma} \cdot \mathbf{p}) = p^2 \tag{1.36}$$

where $\boldsymbol{\sigma}$ are the Pauli spin matrices [52]

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.37}$$

Eq.(1.36) is derived from the more general relation

$$(\boldsymbol{\sigma} \cdot \mathbf{P})\,(\boldsymbol{\sigma} \cdot \mathbf{Q}) = \mathbf{P} \cdot \mathbf{Q} + i\boldsymbol{\sigma} \cdot (\mathbf{P} \times \mathbf{Q}) \tag{1.38}$$

This is an important relation apearing in many different contexts. It has the nice property that it extracts a spin-independent termf from a spin-dependent operator expression.

---

[5] The phase indeterminacy is closely related to the gauge(phase) invariance Eq.(1.24) introduced in the previous section. A phase can be global or local. In the latter case it is a function of space and time coordinates. The physical properties of a free particle is invariant to any global phase change in its wave function. If we require local gauge(phase) invariance, external fields must be introduced [49]

We may play the same ”game” in the non-relativistic domain. We then immediately see that spin is hidden in the kinetic energy operator. We may introduce an external field by the substitution

$$\mathbf{p} \rightarrow \boldsymbol{\pi} = \mathbf{p} + e\mathbf{A}; \qquad E \rightarrow E + e\phi \tag{1.39}$$

Using Eq.(1.38) and the relation

$$\boldsymbol{\pi} \times \boldsymbol{\pi} = -i\boldsymbol{\nabla} \times \mathbf{A} = -i\mathbf{B} \tag{1.40}$$

where ϕ and $\mathbf{A}$ are the scalar and vector potentials respectively, we see that an extra term appears in the non-relativistic Hamiltonian

$$\hat{H} = \frac{\pi^2}{2m} - e\phi + \frac{e}{2m}\left(\boldsymbol{\sigma} \cdot \mathbf{B}\right) \tag{1.41}$$

The term is identified as the spin-Zeeman term and represents the interaction of spin and an external magnetic field $\mathbf{B}$, from which the anomalous Zeeman effect arise.

Continuing the ”game”, it is even possible to derive a four-component non-relativistic equation for spin-$\frac{1}{2}$ particles, as done by Lèvy-Leblond [53]

$$\begin{bmatrix} E & (\boldsymbol{\sigma} \cdot \mathbf{p}) \\ (\boldsymbol{\sigma} \cdot \mathbf{p}) & 2m \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = 0 \tag{1.42}$$

By elimination of the component ϕ_2 we recover Eq.(1.34). What is particularly interesting about Eq.(1.42) is that it is not derived from correspondence principles, but from exploring projective (ray) representations of the non-relativistic analogue of the Lorentz group, the Galilei group. This demonstrates that spin arises naturally in the study of space-time symmetries in both the non-relativistic and relativistic domains. A lucid discussion of this point is given in [54, 55]. We return to the Lévy-Leblond equation in section 1.8.

1.3 The Dirac equation

> *N.Bohr:* – *What are you working on?*
> *P.A.M.Dirac:* – *I'm trying to get a relativistic theory of the electron.*
> *N.Bohr:* – *But Klein has already solved that problem.*
>
> SOLVAY CONFERENCE *(1927) [56]*

Let us now turn to the derivation of the relativistic wave equation for the electron. It will turn out that it is not possible to construct such an equation for the electron alone. Rather we obtainan equation that describes both the electron and its *antiparticle*, the positron.

Straightforward operator substitution Eq.(1.35) in the relativistic free-particle expression Eq.(1.21) leads to the *Klein-Gordon equation*

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\mathbf{p}^2\right]\phi=-\left(mc\right)^2\phi \tag{1.43}$$

The energy operator appears squared in this equation. Hence it has solutions of both positive and negative energies. Contrary to the classical case, the negative energy solutions cannot be discarded since our functional space would then become incomplete. There will always be a finite probability of transitions between states of negative and positive energies. Another problem is that the negative energy solutions have negative probability densities. These difficulties led to the rejection of the Klein-Gordon equation[6]. Instead, Dirac tried to linearize the energy expression Eq.(1.21) by exploiting Eq.(1.36). "It took me quite a while . . . before I suddenly realized that there was no need to stick to quantities $\boldsymbol{\sigma}$. . . with just two rows and colums. Why not go to four rows and columns?" [56]. This lead to the introduction of the Dirac $\boldsymbol{\alpha}$ and β matrices

$$\boldsymbol{\alpha}=\left[\begin{array}{cc}0&\boldsymbol{\sigma}\\\boldsymbol{\sigma}&0\end{array}\right];\qquad\beta=\left[\begin{array}{cc}I_2&0\\0&-I_2\end{array}\right];\qquad\left[\alpha_q,\beta\right]_+=0,\quad q=x,y,z \tag{1.44}$$

and the Dirac equation [58, 59].

We shall derive the Dirac equation following an approach introduced by van Waerden [60]. We expand the scalar wave function Eq.(1.43) using Pauli matrices

$$-\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\mathbf{p}^2\right]\phi=\left[\frac{i}{c}\frac{\partial}{\partial t}-(\boldsymbol{\sigma}\cdot\mathbf{p})\right]\left[\frac{i}{c}\frac{\partial}{\partial t}+(\boldsymbol{\sigma}\cdot\mathbf{p})\right]\phi=\left(mc\right)^2\phi \tag{1.45}$$

where ϕ is a two-component wave function. To obtain a first-order equation we introduce

$$\phi_1=\frac{1}{mc}\left[\frac{i}{c}\frac{\partial}{\partial t}+(\boldsymbol{\sigma}\cdot\mathbf{p})\right]\phi;\qquad\phi_2=\phi \tag{1.46}$$

---

[6] The Klein-Gordon equation was revived in 1934 as the relativistic wave equation for spinless charged particles [57].

The second-order equation then becomes equivalent to two first-order equations

$$\begin{aligned}\left[\frac{i}{c}\frac{\partial}{\partial t}-(\boldsymbol{\sigma}\cdot\mathbf{p})\right]\phi_1 &= mc\phi_2 \qquad (a)\\ \left[\frac{i}{c}\frac{\partial}{\partial t}+(\boldsymbol{\sigma}\cdot\mathbf{p})\right]\phi_2 &= mc\phi_1 \qquad (b)\end{aligned} \tag{1.47}$$

To relate this to the Dirac equation for the free electron we take sums and differences

$$\begin{aligned}\frac{i}{c}\frac{\partial}{\partial t}\left[\phi_1+\phi_2\right]-(\boldsymbol{\sigma}\cdot\mathbf{p})\left[\phi_1-\phi_2\right] &= mc\left[\phi_1+\phi_2\right] \qquad (a+b)\\ (\boldsymbol{\sigma}\cdot\mathbf{p})\left[\phi_1+\phi_2\right]-\frac{i}{c}\frac{\partial}{\partial t}\left[\phi_1+\phi_2\right] &= mc\left[\phi_1-\phi_2\right] \qquad (b-a)\end{aligned} \tag{1.48}$$

and introduce the notation

$$\psi^L=\phi_1+\phi_2;\quad \psi^S=\phi_1-\phi_2 \tag{1.49}$$

We then obtain

$$\left[\begin{array}{cc}\frac{i}{c}\frac{\partial}{\partial t} & -(\boldsymbol{\sigma}\cdot\mathbf{p})\\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -\frac{i}{c}\frac{\partial}{\partial t}\end{array}\right]\left[\begin{array}{c}\psi^L\\ \psi^S\end{array}\right]=mc\left[\begin{array}{c}\psi^L\\ \psi^S\end{array}\right] \tag{1.50}$$

The 4-component can be completely in terms of 4-vectors and scalar quantities as

$$\left(i\gamma_\mu\partial_\mu-mc\right)\psi=0;\qquad \gamma_\mu=\left(\beta\boldsymbol{\alpha},i\beta\right),\quad \psi=\left[\begin{array}{c}\psi^L\\ \psi^S\end{array}\right] \tag{1.51}$$

and is therefore manifestly Lorentz invariant Dirac. The Dirac equation in its more familiar form is straightforwardly obtained by multiplication with βc from the left

$$\left[i\frac{\partial}{\partial t}-c\left(\boldsymbol{\alpha}\cdot\mathbf{p}\right)\right]\psi=\beta mc^2\psi \tag{1.52}$$

External fields are introduced by means of the substitutions in Eq.(1.22). The Dirac equation then attains the form

$$\hat{\mathcal{D}}\psi=\left[\hat{h}_D-i\frac{\partial}{\partial t}\right]\psi=0;\qquad h_D=\beta mc^2+c\left(\boldsymbol{\alpha}\cdot\boldsymbol{\pi}\right)-e\phi \tag{1.53}$$

or, in full,

$$\begin{bmatrix} mc^2 - e\phi - i\frac{\partial}{\partial t} & 0 & ecA_z - icd_z & ecA_- - icd_- \\ 0 & mc^2 - e\phi - i\frac{\partial}{\partial t} & ecA_+ - icd_+ & -ecA_z + icd_z \\ ecA_z - icd_z & ecA_- - icd_- & -mc^2 - e\phi - i\frac{\partial}{\partial t} & 0 \\ ecA_+ - icd_+ & -ecA_z - icd_z & 0 & -mc^2 - e\phi - i\frac{\partial}{\partial t} \end{bmatrix} \begin{bmatrix} \psi^{L\alpha} \\ \psi^{L\beta} \\ \psi^{S\alpha} \\ \psi^{S\beta} \end{bmatrix} = 0 \tag{1.54}$$

where we have introduced the notation

$$d_z = \frac{\partial}{\partial z}; \qquad d_\pm = \frac{\partial}{\partial x} \pm i\frac{\partial}{\partial y}. \tag{1.55}$$

The relativistic wave equation is seen to have four components. The extra degrees of freedom results from the introduction of spin and the fact that the equation describes both the electron and its antiparticle, the positron. The upper two components are dominant in electronic solutions and are therefore denoted the large (L) components, whereas the two lower components are denoted the small components. The large and small components both have a spin-up (α) and a spin-down (β) part. In the next three sections we will explore the physical content of the Dirac equation from its symmetry properties.

1.4 Spatial symmetry

> *One day, while we were walking on the beach, he [Dirac] told me that he would teach me a saying:'It is easy, if you remember the symmetry. Watch the symmetry.' He went on, 'When a man says yes, he means perhaps; when he says perhaps, he means no; when he says no, he is no diplomat. When a lady says no, she means perhaps; when she says perhaps, she means yes; when she says yes, she is no lady'. With a couple of repetitions, I learned it, and he was pleased.*
>
> S.A. KURSUNOGLU *(1987) [61]*

A symmetry operation $\hat{G}$ is defined as an operation that commutes with the Hamiltonian of the system:

$$\left[\hat{G}, \hat{H}\right] = 0 \tag{1.56}$$

Symmetry operations are either unitary or antiunitary, as shown by the following argument by Wigner [62]: Observables calculated from a given wave function are invariant under any symmetry operation on the wave function. For the transition probability Eq.(1.33) we must therefore have

$$\left|\left\langle \hat{G}\Psi_i \mid \hat{G}\Psi_j \right\rangle\right|^2 = \left|\left\langle \Psi_i \mid \Psi_j \right\rangle\right|^2 = \left\langle \Psi_i \mid \Psi_j \right\rangle \left\langle \Psi_i \mid \Psi_j \right\rangle^* \tag{1.57}$$

where the interaction operator $\hat{\Omega}$ has been set equal to one for simplicity (to obtain a totally symmetric operator). The above relation can be realized by

$$\begin{array}{rcll} \left\langle \hat{G}\Psi_i \mid \hat{G}\Psi_j \right\rangle & = & \left\langle \Psi_i \mid \Psi_j \right\rangle; & \Rightarrow \quad \hat{G} \text{ is unitary} \\ \left\langle \hat{G}\Psi_i \mid \hat{G}\Psi_j \right\rangle & = & \left\langle \Psi_j \mid \Psi_i \right\rangle; & \Rightarrow \quad \hat{G} \text{ is anti-unitary} \end{array} \tag{1.58}$$

Let us first consider unitary symmetry operations and defer the discussion of antiunitary operators until the next section. A more thorough discussion will be found in Appendix C. Here we will exploit a simple, but powerful observation: In the absence of any external field the Dirac Hamiltonian Eq.(1.53) must be invariant under all possible symmetry operations (unitary or antiunitary) of time and space. This follows from the homogeneity of space and time and from the isotropy of space. The latter implies rotational invariance and the conservation of total angular momentum. The Dirac Hamiltonian does not, however, commute with the orbital angular mometum operator $\mathbf{l}$, which means that some angular momentum is "missing". By inspection, we find that the Dirac Hamiltonian commutes with a total angular momentum $\mathbf{j} = \mathbf{l} + \mathbf{s}$, which demonstrates that the Dirac equation describes particles of spin $s = \frac{1}{2}$. The spin-operator $\mathbf{s}$ is represented by $\frac{1}{2}\rho\boldsymbol{\alpha}$ where $\rho\boldsymbol{\alpha}$ is the 4×4 analogues of the Pauli spin matrices

$$\rho\boldsymbol{\alpha} = (I_2 \otimes \boldsymbol{\sigma}) = \left[\begin{array}{cc} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{array}\right]; \qquad \rho = \left[\begin{array}{cc} 0 & I_2 \\ I_2 & 0 \end{array}\right]; \quad \begin{array}{rcl} [\rho, \beta]_+ & = & 0 \\ [\rho, \boldsymbol{\alpha}] & = & 0 \end{array} \tag{1.59}$$

The total angular momentum operator $\mathbf{j}$ is the generator of infinitesimal rotations. A finite rotation ϕ about an axis represented by the unit vector $\mathbf{n}$ is given by

$$R(\phi, \mathbf{n}) = e^{-i\phi(\mathbf{n}\cdot\mathbf{j})} = e^{-i\phi(\mathbf{n}\cdot\mathbf{l})}e^{-i\phi(\mathbf{n}\cdot\mathbf{s})} = R_r(\phi, \mathbf{n})\, R_\sigma(\phi, \mathbf{n}) \qquad [\mathbf{l}, \mathbf{s}] = 0 \tag{1.60}$$

Note from the above relation that the rotation operator splits into one part acting on spatial coordinates and one part acting on spin coordinates. The rotation operator for the spin part is straightforwardly established as (see Appendix C)

$$R_\sigma(\phi, \mathbf{n}) = \cos\frac{1}{2}\phi - i(\rho\boldsymbol{\alpha}\cdot\mathbf{n})\sin\frac{1}{2}\phi \tag{1.61}$$

In particular binary rotations about main axes are given by

$$C_2^x = -i\rho\boldsymbol{\alpha}_x; \quad C_2^y = -i\rho\boldsymbol{\alpha}_y; \quad C_2^z = -i\rho\boldsymbol{\alpha}_z \tag{1.62}$$

We note that $C_2^q C_2^q = -I_4$ for all coordinates q which demonstrates that fermion functions change sign upon a rotation 2π, in contrast to boson functions for which a rotation 2π is equivalent to the identity operation. The fermion phase shift has been verified experimentally in neutron [63, 64, 65] and NMR [66] interferometry.

Representations in spin space of other symmetry operations can be derived using the fact that the operator $(\boldsymbol{\alpha}\cdot\mathbf{p})$ must be invariant under all symmetry operations. Hence the Dirac $\boldsymbol{\alpha}$-matrices must transform as the momentum operator $\mathbf{p}$, that is, as the Cartesian coordinates, so that we have

$$\hat{G}_\sigma \boldsymbol{\alpha} \hat{G}_\sigma^{-1} = \hat{G}_r \mathbf{p} \hat{G}_r^{-1} \tag{1.63}$$

where $\hat{G}_r$ and $\hat{G}_\sigma$ act on spatial and spin coordinates, respectively. Rotations in spin space and ordinary space are connected through Eq.(1.63), which establishes a mapping between the group of SU(2) of special unitary matrices (acting on 2-spinors) and the group SO(3) of special orthogonal matrices (acting on Cartesian coordinates). However, the mapping is a two-to-one homomorphism since a rotation 2π is equivalent to the identity operation in ordinary space, whereas it leads to a phase shift in spin space. Note that the above relations leave a complex phase undetermined for $\hat{G}_\sigma$, in accordance with the phase indeterminacy of the wave function.

Under space inversion the momentum operator changes sign, so that the spin part of the space inversion operator is determined by

$$\hat{i}_\sigma \boldsymbol{\alpha} \hat{i}_\sigma^{-1} = -\boldsymbol{\alpha} \tag{1.64}$$

The above relation implies that the inversion operator $\hat{i}_\sigma$ anticommutes with the Dirac $\boldsymbol{\alpha}$-matrices. An obvious choice for the inversion operator is therefore the Dirac β-matrix. We define

$$\hat{i}_\sigma = -i\beta \tag{1.65}$$

The fact that the Dirac β-matrix appears in the spin part of the inversion operator demonstrates that the large ψ^L and small ψ^S components have opposite parity[7] Finally we derive expressions for the operations of reflection in the spin coordinates using the fact that reflections are the product of inversion and binary rotations :

$$\hat{\sigma}_{yz} = -\beta\rho\boldsymbol{\alpha}_x; \quad \hat{\sigma}_{zx} = -\beta\rho\boldsymbol{\alpha}_y; \quad \hat{\sigma}_{xy} = -\beta\rho\boldsymbol{\alpha}_z \tag{1.66}$$

We arrive at the same forms (within a phase factor) using Eq.(1.63), for example

$$\hat{\sigma}_{yz}\left(\alpha_x, \alpha_y, \alpha_z\right)\hat{\sigma}_{yz}^{-1} = \left(-\alpha_x, \alpha_y, \alpha_z\right) \quad \Rightarrow \quad \hat{\sigma}_{yz} = -\beta\rho\boldsymbol{\alpha}_x \tag{1.67}$$

Two-component analogues of the symmetry operations derived so far are obtained by the substitutions

$$\rho\boldsymbol{\alpha} \to \boldsymbol{\sigma}; \quad \to \quad \beta \to I_2 \tag{1.68}$$

A problematic aspect of the representations derived so far is that they are at odds with the conventions of double group theory. Double groups were introduced as an artifice by Bethe [69] to avoid two-valued representations of fermion functions, which are not true representations of the symmetry group, and thereby recover the whole machinery of group theory. He introduced an extra element $\overline{E}$, corresponding to a rotation 2π about an arbitrary axis. This leads to a doubling of the number of symmetry operations of the group, but generally not to a doubling of the number of irreducible representations (irreps). The extra irreps that appear in the double groups are spanned by fermion functions and are consequently denoted fermion irreps, whereas the irreps of the corresponding single groups are boson irreps.

---

[7] Note from Eq.(1.49) that the functions ϕ_1 and ϕ_2 are not eigenfunctions of parity. Rather, the parity operator takes ϕ_1 into ϕ_2 and vice versa. The Eqs.(1.47) decouple for rest mass m = 0. Eq.(1.47b) was therefore proposed as the wave equation for a massless spin-$\frac{1}{2}$ particle in 1929 by Weyl [67], but was rejected since the wave function ϕ_1 is not invariant under parity. Parity is, however, not conserved in weak interactions. With the demonstration in 1957 of the violation of parity conservation in the β-decay of the ^{60}Co - nucleus [68], the Weyl equation was revived as a two-component equation for the neutrino.

The introduction of $\overline{E}$ doubles the order of all rotations. On the other hand, the order of inversion is still taken to be two ($\hat{i}^2 = E$), since space inversion commutes with all rotations in ordinary space. In our representation, Eq.(1.65), however, inversion is of order four. Since there is an inherent phase indeterminacy, we could correct this by changing the phase of our inversion operator. However, we can show that operation of inversion *must* be of order four if we want consistency in the representation of spatial symmetry operations.

It is well known that two spins $s = \frac{1}{2}$ couple to a singlet function and the three components of a triplet function. The latter three functions transform as the spherical harmonics Y_{lm} with $l = 1$. By forming the direct product of a 2×2 matrix representation of a symmetry operation (in the two-component case) with itself, we obtain a link to matrix representations of the spherical harmonics with l = 1. The direct product gives a 4×4 matrix from which we by a unitary transformation can isolate a 3×3 block representing the corresponding symmetry operation in the basis of spherical harmonics for l = 1. The phases for the symmetry operations presented above have been chosen with care so as to obtain agreement with the Condon-Shortley phase convention for spherical harmonics [70]. In the case of inversion, the case is, however, unambiguous: The direct product of the two-component representation of inversion with itself gives

$$-iI_2 \otimes -iI_2 = -I_4 \tag{1.69}$$

The identity matrix is invariant under all unitary transformations, and so we obtain $-I_3$ as the representation of inversion in the basis of $\{Y_{1,1,}, Y_{1,0}, Y_{1,-1}\}$, as we should. It is not possible to obtain the same representation starting from a two-component inversion operator of order only two. Altmann [71], in the language of projective representations, sees this discrepancy between representations merely as a choice of gauge (phase). In my opinion his explanation seems somewhat *ad hoc*. There is a fundamental weakness in the derivation of the behaviour of inversion in double group theory. The extra element $\overline{E}$ is introduced to account for the fact that fermion functions have a behaviour under rotation that is different from rotations in ordinary space. Yet the behaviour of inversion in double group theory is deduced with explicit reference to inversion in ordinary space, which is somewhat inconsistent. It would be interesting to see whether the behaviour of fermion functions under the operations of inversion or reflections can be resolved experimentally.

1.5 Time reversal symmetry and quaternion algebra

> *The invention of the calculus of quaternions is a step towards the knowledge of quantities related to space which can only be compared for its importance, with the invention of triple coordinates by Descartes. The ideas of this calculus, as distinguished from its operations and symbols, are fitted to be of the greatest use in all parts of science.*
>
> J.C.MAXWELL *(1869)*
>
> *Quaternions came from Hamilton after his really good work had been done; and, though beautiful ingenious, have been an unmixed evil to those who have touched them in any way, including Clerk Maxwell.*
>
> LORD KELVIN *(1892) [72]*

We now turn our attention to antiunitary operators. From the previous section we recall that an antiunitary operator $\hat{\mathcal{K}}$ is defined by

$$\left\langle \hat{\mathcal{K}}\phi_1 \mid \hat{\mathcal{K}}\phi_2 \right\rangle = \langle \phi_2 \mid \phi_1 \rangle = \langle \phi_1 \mid \phi_2 \rangle^* = \hat{\mathcal{K}} \langle \phi_1 \mid \phi_2 \rangle \tag{1.70}$$

The last two terms indicate the antilinearity of antiunitary operators:

$$\hat{\mathcal{K}} \left(a\phi_1 + b\phi_2 \right) = a^* \hat{\mathcal{K}}\phi_1 + b^* \hat{\mathcal{K}}\phi_2 \tag{1.71}$$

It is straightforwardly shown that the product of two antiunitary operators is a unitary operator, which implies that any antiunitary operator can be written as a product of a unitary operator and some antiunitary operator. The simplest choice of an operator to fullfill conditions Eq.(1.70) and Eq.(1.71) is the complex conjugation operator $\hat{\mathcal{K}}_0$. A general antiunitary operator may therefore be written as

$$\hat{\mathcal{K}} = U\hat{\mathcal{K}}_0 \tag{1.72}$$

In non-relativistic systems $\hat{\mathcal{K}}_0$ commutes with the Hamiltonian in the absence of external magnetic fields and represents the operation of time reversal [73]. This is straightforwardly seen by letting $\hat{\mathcal{K}}_0$ operate on both sides of the time-dependent Schrödinger equation

$$\begin{array}{rcll}
\hat{\mathcal{K}}_0 \left[i\dfrac{\partial}{\partial t}\Psi\left(\mathbf{r},t\right) \right] & = & \hat{\mathcal{K}}_0 \left[H\Psi\left(\mathbf{r},t\right) \right] & \\
 & \Downarrow & & \\
-i\dfrac{\partial}{\partial t}\hat{\mathcal{K}}_0\Psi\left(\mathbf{r},t\right) = i\dfrac{\partial}{\partial(-t)}\hat{\mathcal{K}}_0\Psi\left(\mathbf{r},t\right) & = & H\hat{\mathcal{K}}_0\Psi\left(\mathbf{r},t\right) & \left[H,\hat{\mathcal{K}}_0\right] = 0 \\
 & \Downarrow & & \\
i\dfrac{\partial}{\partial t}\left(\hat{\mathcal{K}}_0\Psi\left(\mathbf{r},-t\right)\right) & = & H\left(\hat{\mathcal{K}}_0\Psi\left(\mathbf{r},-t\right)\right) &
\end{array} \tag{1.73}$$

We consider next the form of the time reversal operator $\hat{\mathcal{K}}$ in the 4-component formalism. Momenta are reversed under the operation of time reversal

$$\hat{\mathcal{K}}\mathbf{p}\hat{\mathcal{K}}^{-1} = -\mathbf{p} \tag{1.74}$$

Using the general form Eq.(1.72) of an antiunitary operator and Eq.(1.63) we therefore find that

$$\hat{\mathcal{K}}\boldsymbol{\alpha}\hat{\mathcal{K}}^{-1} = U\boldsymbol{\alpha}^* U^{-1} = -\boldsymbol{\alpha} \quad \Rightarrow \quad U\left(-\alpha_x, \alpha_y, -\alpha_z\right)U^{-1} = \left(\alpha_x, \alpha_y, \alpha_z\right) \tag{1.75}$$

Since the Dirac $\boldsymbol{\alpha}$ matrices transform as the coordinates, we identify U as a rotation about the y-axis Eq.(1.62) and write the time reversal operator as

$$\hat{\mathcal{K}} = -i\rho\alpha_y\hat{\mathcal{K}}_0 = -i\left(I_2 \otimes \sigma_y\right)\hat{\mathcal{K}}_0 \tag{1.76}$$

Since we are mainly interested in fermion functions, we can alternatively define the time reversal operator by its action on a fermion function ϕ, that is

$$\hat{\mathcal{K}}a\phi = a^*\hat{\mathcal{K}}\phi; \qquad \hat{\mathcal{K}}^2\phi = -\phi \tag{1.77}$$

We shall use the convention

$$\hat{\mathcal{K}}\phi = \overline{\phi} \tag{1.78}$$

and denote ϕ and $\overline{\phi}$ as *Kramers partners.* We shall now use the alternative definition Eq.(1.77) to derive the general matrix structure of Hermitian operators $\hat{\Omega}_\pm$ that are symmetric(+) or antisymmetric(−) under time reversal [33, 74]

$$\hat{\mathcal{K}}\hat{\Omega}_t\hat{\mathcal{K}}^{-1} = t\hat{\Omega}_t; \qquad t = \pm 1 \tag{1.79}$$

We consider the matrix representation of $\hat{\Omega}_t$ in a *Kramers restricted basis* which we define as follows: Operate with $\hat{\mathcal{K}}$ on a set of fermion basis functions $\{\phi_p\}$ to generate a complementary basis $\left\{\overline{\phi}_p\right\}$. The Kramers restricted basis is then union of the two sets. We establish the following relations between matrix elements of $\hat{\Omega}_t$

$$\begin{array}{rcccccl} \Omega_{\overline{pq}} & = & ca\hat{l}K\Omega^*_{\overline{pq}} & = & \left\langle \hat{\mathcal{K}}\overline{\phi}_q \middle| \hat{\mathcal{K}}\Omega\hat{\mathcal{K}}^{-1} \middle| \hat{\mathcal{K}}\overline{\phi}_p \right\rangle & = & t\Omega^*_{pq} \\ \Omega_{\overline{p}q} & = & ca\hat{l}K\Omega^*_{\overline{p}q} & = & \left\langle \hat{\mathcal{K}}\overline{\phi}_p \middle| \hat{\mathcal{K}}\Omega\hat{\mathcal{K}}^{-1} \middle| \hat{\mathcal{K}}\phi_q \right\rangle & = & -t\Omega^*_{p\overline{q}} \end{array} \tag{1.80}$$

From these relations we find that the matrix representation of $\hat{\Omega}_\pm$ has the structure

$$\boldsymbol{\Omega}_t = \left[\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ -t\mathbf{B}^* & t\mathbf{A}^* \end{array}\right]; \qquad \begin{array}{rclrcl} \mathbf{A}^\dagger & = & \mathbf{A}; & A_{pq} & = & \Omega_{pq} \\ \mathbf{B}^T & = & -\mathbf{B}; & B_{pq} & = & \Omega_{p\overline{q}} \end{array} \tag{1.81}$$

Let us investigate the properties of the above matrix structure. Since $\Omega_{\pm}$ is Hermitian, its matrix may be diagonalized by a unitary transformation, giving real eigenvalues ε :

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -t\mathbf{B}^{*} & t\mathbf{A}^{*} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{\alpha} \\ \mathbf{c}^{\beta} \end{bmatrix} = \varepsilon \begin{bmatrix} \mathbf{c}^{\alpha} \\ \mathbf{c}^{\beta} \end{bmatrix} \tag{1.82}$$

We write out the corresponding matrix equations

$$\begin{array}{rcrcl} \mathbf{A}\mathbf{c}^{\alpha} & + & \mathbf{B}\mathbf{c}^{\beta} & = & \varepsilon\mathbf{c}^{\alpha} \\ -t\mathbf{B}^{*}\mathbf{c}^{\alpha} & + & t\mathbf{A}^{*}\mathbf{c}^{\beta} & = & \varepsilon\mathbf{c}^{\beta} \end{array} , \tag{1.83}$$

conjugate both equations and then multiply the first with t and the second with $-t$. This gives

$$\begin{array}{rcrcr} t\mathbf{A}^{*}\mathbf{c}^{\alpha *} & + & t\mathbf{B}^{*}\mathbf{c}^{\beta *} & = & t\varepsilon\mathbf{c}^{\alpha *} \\ \mathbf{B}\mathbf{c}^{\alpha *} & - & \mathbf{A}\mathbf{c}^{\beta *} & = & -t\varepsilon\mathbf{c}^{\beta *} \end{array} \tag{1.84}$$

which can be expressed on matrix form as

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -t\mathbf{B}^{*} & t\mathbf{A}^{*} \end{bmatrix} \begin{bmatrix} -\mathbf{c}^{\beta *} \\ \mathbf{c}^{\alpha *} \end{bmatrix} = t\varepsilon \begin{bmatrix} -\mathbf{c}^{\beta *} \\ \mathbf{c}^{\alpha *} \end{bmatrix} \tag{1.85}$$

Hence we can conclude the following about the matrix of $\hat{\Omega}$ in a Kramers restricted basis

- If $\hat{\Omega}$ is *symmetric* with respect to time reversal, its matrix is doubly degenerate with eigenvectors related by time reversal symmetry

  $$\left\{ \begin{bmatrix} \mathbf{c}^{\alpha} \\ \mathbf{c}^{\beta *} \end{bmatrix}, \begin{bmatrix} -\mathbf{c}^{\beta} \\ \mathbf{c}^{\alpha *} \end{bmatrix} \right\} \tag{1.86}$$

  To some extent time reversal symmetry recovers the spin symmetry lost in the relativistic domain, but the recovery is only partial. In the non-relativistic domain a totally symmetric (spinfree) operator does not couple two spin orbitals if they have opposite spin. In the relativistic domain we have the weaker relation

  $$\left\langle \phi_i \left| \hat{\Omega}_{+} \right| \overline{\phi}_j \right\rangle \equiv 0 \quad \text{only if } i = j \tag{1.87}$$

- If $\hat{\Omega}$ is *antisymmetric* with respect to time reversal, then eigenvectors are pairwise related by time reversal symmetry Eq.(1.86) that have eigenvalues of the same absolute value, but with opposite signs.

Let us now investigate time reversal symmetry in the Dirac equation. It turns out that this is best done by a reordering of the 4-spinors:

$$\begin{bmatrix} \psi^L \\ \psi^S \end{bmatrix} = \begin{bmatrix} \psi^{L\alpha} \\ \psi^{L\beta} \\ \psi^{S\alpha} \\ \psi^{S\beta} \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \psi^{L\alpha} \\ \psi^{S\alpha} \\ \psi^{L\beta} \\ \psi^{S\beta} \end{bmatrix} = \begin{bmatrix} \psi^{\alpha} \\ \psi^{\beta} \end{bmatrix} \tag{1.88}$$

From the matrix structures derived above we can immediately split the Dirac equation Eq.(1.53) into one part that is symmetric and another part that is antisymmetric under the operation of time reversal

$$\hat{\mathcal{D}}\psi = \left[\hat{\mathcal{D}}^+ + \hat{\mathcal{D}}^-\right]\psi = 0 \tag{1.89}$$

The symmetric part is

$$\hat{\mathcal{D}}^+ = \begin{bmatrix} mc^2 - e\phi & -icd_z & 0 & -icd_- \\ -icd_z & -mc^2 - e\phi & -icd_- & 0 \\ 0 & -icd_+ & mc^2 - e\phi & icd_z \\ -icd_+ & 0 & -icd_z & -mc^2 - e\phi \end{bmatrix} \tag{1.90}$$

and the antisymmetric part is

$$\hat{\mathcal{D}}^- = \begin{bmatrix} -i\frac{\partial}{\partial t} & ecA_z & 0 & ecA_- \\ ecA_z & -i\frac{\partial}{\partial t} & ecA_- & 0 \\ 0 & ecA_+ & -i\frac{\partial}{\partial t} & -ecA_z \\ ecA_+ & 0 & -ecA_z & -i\frac{\partial}{\partial t} \end{bmatrix} \tag{1.91}$$

We can now explicitly show that the pair of eigenvectors in Eq.(1.86) are related by time reversal symmetry. With reordered spinors (Eq.1.88) the time reversal operator has the form

$$\hat{\mathcal{K}} = -\mathrm{i}\left[\sigma_y \otimes I_2\right]\hat{\mathcal{K}}_0 \tag{1.92}$$

Operating with $\hat{\mathcal{K}}$ on an eigenvector $\mathbf{c}$ we obtain

$$\hat{\mathcal{K}}\mathbf{c} = \begin{bmatrix} 0 & -I_2 \\ I_2 & 0 \end{bmatrix} \hat{\mathcal{K}}_0 \begin{bmatrix} \mathbf{c}^{\alpha} \\ \mathbf{c}^{\beta} \end{bmatrix} = \begin{bmatrix} -\mathbf{c}^{\beta *} \\ \mathbf{c}^{\alpha *} \end{bmatrix} = \overline{\mathbf{c}} \tag{1.93}$$

as required.

By restricting the Dirac operator to the time symmetric part $\hat{\mathcal{D}}^{+}$ only, a considerable simplification is possible by the introduction of quaternion algebra. A (real) quaternion number[8] is written as

$$q = \sum_{\Lambda=0}^{3} c_{\Lambda} e_{\Lambda} = c_0 + c_1 \check{\imath} + c_2 \check{\jmath} + c_3 \check{\mathrm{k}}; \quad c_{\Lambda} \in \grave{\mathrm{R}} \tag{1.94}$$

in which the quaternion units $\check{\imath}$, $\check{\jmath}$, and $\check{\mathrm{k}}$ obey the following multiplication rules

$$\check{\imath}^2 = \check{\jmath}^2 = \check{\mathrm{k}}^2 = \check{\imath}\check{\jmath}\check{\mathrm{k}} = -1 \tag{1.95}$$

The quaternion units are equivalent in the sense that they may be interchanged by cyclic permutation $\check{\imath} \rightarrow \check{\jmath} \rightarrow \check{\mathrm{k}} \rightarrow \check{\imath}$. Thus, in a complex number $a + \mathrm{i}b$ the imaginary i may correspond to either $\check{\imath}$, $\check{\jmath}$, or $\check{\mathrm{k}}$ without changing its algebraic properties.

When Pauli introduced the spin matrices that bear his name, Jordan pointed out[76] that the properties of imaginary i times the Pauli matrices were identical to that of the quaternion units $\check{\imath}$,$\check{\jmath}$ and $\check{\mathrm{k}}$. Specifically we have the mapping

$$i\sigma_z \leftrightarrow \check{\imath}; \qquad i\sigma_y \leftrightarrow \check{\jmath}; \qquad i\sigma_x \leftrightarrow \check{\mathrm{k}} \tag{1.96}$$

which allow us to represent a quaternion number by a 2 × 2 complex matrix

$$q = a + b\check{\jmath} \leftrightarrow Q = c_0 I_2 + c_1 i\sigma_z + c_2 i\sigma_y + c_3 i\sigma_x = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}; \qquad \begin{array}{l} a = c_0 + ic_1; \\ b = c_2 + ic_3; \end{array} \tag{1.97}$$

so that

$$q_1 q_2 \leftrightarrow Q_1 Q_2 \tag{1.98}$$

This is analogous to the complex numbers, which may be represented by 2×2 real matrices. Two equivalent representations exist

$$c = a + \mathrm{i}b \leftrightarrow \left\{ \begin{array}{lcl} C & = & \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \\ & & \\ C' & = & \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \end{array} \right. \qquad a, b, \in \grave{\mathrm{R}} \tag{1.99}$$

---

[8] Note that quaternion numbers are not *quaternionic*, just as complex numbers are not *complexionic*. The use of quaternion algebra in physics is described in [45, 75].

The quaternion analogue of C’ would be

$$Q' = \begin{bmatrix} a & -b^* \\ b & a^* \end{bmatrix} \tag{1.100}$$

However we find that $Q_1'Q_2' \not\leftrightarrow q_1q_2$. Instead we have $Q_1'Q_2' \leftrightarrow q_2q_1$. This demonstrates a troublesome feature of quaternion numbers, namely that they do not commute under multiplication[9].

Quaternion algebra allows block diagonalization of the matrix of an operator that is symmetric under time reversal. By comparing Eq.(1.81) and Eq.(1.97) we see that the matrix $\mathbf{\Omega}_+$ has a structure identical to that of the 2×2 complex matrix representation Q of quaternion numbers, which means that it can be expressed in terms of Pauli matrices or quaternion units. Block diagonalization is achieved through the unitary quaternion transformation

$$\mathbf{U}^\dagger \mathbf{\Omega}_+ \mathbf{U} = \begin{bmatrix} \mathbf{A} + \mathbf{B}\check{\jmath} & 0 \\ 0 & -\check{k}\,(\mathbf{A} + \mathbf{B}\check{\jmath})\,\check{k} \end{bmatrix}; \qquad \mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I} & \check{\jmath}\mathbf{I} \\ \check{\jmath}\mathbf{I} & \mathbf{I} \end{bmatrix} \tag{1.101}$$

Due to the decoupling of blocks, the transformation leads to an exact reduction of the time reversal symmetric Dirac operator $\hat{\mathcal{D}}^+$ Eq.(1.90) to two-component form, albeit in terms of quaternion algebra (indicated by upper prescript Q):

$${}^Q\hat{h}\,{}^Q\psi = E\,{}^Q\psi \tag{1.102}$$

where

$${}^Q\hat{h} = \left\{ \begin{bmatrix} mc^2 - e\phi & 0 \\ 0 & -mc^2 - \phi \end{bmatrix} - c\check{\imath} \begin{bmatrix} 0 & d_z \\ d_z & 0 \end{bmatrix} - c\check{\jmath} \begin{bmatrix} 0 & d_y \\ d_y & 0 \end{bmatrix} - c\check{k} \begin{bmatrix} 0 & d_x \\ d_x & 0 \end{bmatrix} \right\} \tag{1.103}$$

The quaternion eigenfunctions ${}^Q\psi$ are related to the corresponding complex reordered 4-spinors Eq.(1.90) by

$${}^Q\psi = \left[\psi^{\alpha *} - \psi^{\beta *}\check{\jmath}\right] \quad \leftrightarrow \quad \begin{bmatrix} \psi^\alpha \\ \psi^\beta \end{bmatrix} \tag{1.104}$$

The quaternion Dirac operator has an intriguing structure. The scalar potential enter the real part, whereas the kinetic energy part is spanned by the quaternion units $\check{\imath}$,$\check{\jmath}$ and $\check{k}$. The equivalence of quaternion units thus parallels the equivalence of the coordinate axes

---

[9] A historical note: In the early days of quantum mechanics Dirac introduced the concept of c-numbers and q-numbers where ”c stands for classical or maybe commuting” and ”q stands for quantum or maybe queer”[56]. Maybe they should stand for complex and quaternion?

(x,y,z). Note that in the quaternion formulation the time reversal operator is mapped into $-\check{\jmath}$, for we have

$$-\check{\jmath}\,{}^{Q}\psi = \left[-\psi^{\beta} - \psi^{\alpha}\check{\jmath}\right] \quad \leftrightarrow \quad \begin{bmatrix} -\psi^{\beta *} \\ \psi^{\alpha *} \end{bmatrix} = \hat{\mathcal{K}} \begin{bmatrix} \psi^{\alpha} \\ \psi^{\beta} \end{bmatrix} \tag{1.105}$$

Before concluding this section, two final remarks should be made:

1. The full Dirac equation Eq.(1.54) may be expressed in terms of *complex quaternions*. Complex quaternions[77, 78] are obtained by replacing the real coefficients in Eq.(1.94) by complex coefficients. The substitutions $\mathbf{A} = i\mathbf{A}'$ and $\mathbf{B} = i\mathbf{B}'$ in Eq.(1.81) establishes the relation $\mathbf{\Omega}_{-} = i\mathbf{\Omega}'_{+}$. In the complex quaternion Dirac equation the time symmetric $\hat{\mathcal{D}}^{+}$ and antisymmetric $\hat{\mathcal{D}}^{-}$ parts enter the real and imaginary parts of the coefficients, respectively. We shall not pursue this approach here and refer to the literature[79, 80, 81, 82, 83] for details.

2. For a unimodular number the matrix C Eq.(1.99) is identical to a 2×2 orthogonal matrix (see Eq.(1.5)), which represents a two-dimensional rotation. The quaternion units $\check{\imath}$, $\check{\jmath}$ and $\check{k}$ form the basis for general rotations in three-dimensional space, a feature that is seen from Eq.(1.62). In fact the eight basic binary spatial symmetry operations can be mapped into the complex quaternion units as shown in Tab.1.5.

Table 1.1: Mapping between the basic eight binary symmetry operations and the complex quaternion units

Proper rotations	Improper rotations
$\hat{E} \rightarrow I_2 \rightarrow 1$	$\hat{i} \rightarrow -iI_2 \rightarrow -\mathrm{i}$
$\hat{C}_2^z \rightarrow -i\sigma_z \rightarrow -\check{\mathrm{i}}$	$\hat{\sigma}_{xy} \rightarrow -\sigma_z \rightarrow \mathrm{i}\check{\mathrm{i}}$
$\hat{C}_2^y \rightarrow -i\sigma_y \rightarrow -\check{\mathrm{j}}$	$\hat{\sigma}_{zx} \rightarrow -\sigma_y \rightarrow \mathrm{i}\check{\mathrm{j}}$
$\hat{C}_2^x \rightarrow -i\sigma_x \rightarrow -\check{\mathrm{k}}$	$\hat{\sigma}_{yz} \rightarrow -\sigma_x \rightarrow \mathrm{i}\check{\mathrm{k}}$

1.6 Charge conjugation

> *Sehr viel unglücklicher bin ich über die Frage nach der relativistischen Formulierung und über die Inkonsequenz der Dirac-Theori ... Also ich find' die gegenwärtige Lage ganz absurd und hab' mich deshalb, quasi aus Verzweiflung, auf ein anderes gebiet, das des Ferromagnetismus begeben.*
>
> W. HEISENBERG *(1928) [61]*

In the previous sections we explored various symmetries of the Dirac equation. In general a symmetry operator, that is an operator which commutes with the Hamiltonian, introduces a degeneracy, where the degenerate solutions are related by the symmetry operation. On the other hand, an "antisymmetric" operator, that is an operator which anticommute with the Hamiltonian, leads to solutions related by the operator and that have energies of the same absolute value, but opposite signs. Such an operator is obviously at play in the free-particle Dirac equation. It has positive and negative energy branches separated by an energy gap of $2mc^2$. Positive and negative eigenvalues are pairwise related, which corresponds to taking the positive or negative root in the corresponding classical free-particle energy expression Eq.(1.21).

The pairing of eigenvalues is, however, lifted with the introduction of an external field. Consider some positive energy solution of the Dirac equation in the presence of an external field

$$\left[\beta mc + (\boldsymbol{\alpha}\cdot(\mathbf{p}+e\mathbf{A})) - \frac{e}{c}\phi - p_0\right]\psi = 0 \tag{1.106}$$

Its classical analogue is obtained by taking the positive square root in the corresponding energy expression Eq.(1.27).

$$E = +\sqrt{m^2c^4 + c^2\left(\mathbf{p}+e\mathbf{A}\right)^2} - e\phi \tag{1.107}$$

In order to arrive at a negative energy of the same absolute value we have not only to choose the negative square root, but reverse the sign of the momentum and the external fields as well. An alternative to sign reversal of the fields is to introduce a positive charge +e. This means that a positive energy solution of the Dirac equation for a particle of charge −e is a negative energy solution of the Dirac equation for a particle of charge +e:

$$\left[\beta mc + (\boldsymbol{\alpha}\cdot(\mathbf{p}-e\mathbf{A})) + \frac{e}{c}\phi - p_0\right]\psi = 0 \tag{1.108}$$

The two Dirac equations are related by the operation of *charge conjugation* $\hat{\mathcal{C}}$, which anticommutes with the Dirac Hamiltonian in the limit of no external field. We shall derive the explicit form of this operator.

We first note that the sign of the momentum may be reversed by complex conjugation. If we perform this operation on the Dirac equation for charge –e and then multiply both sides with a minus sign, we obtain

$$\left[-\beta mc + (\boldsymbol{\alpha}^* \cdot (\mathbf{p} - e\mathbf{A})) + \frac{e}{c}\phi - p_0\right]\psi = 0 \tag{1.109}$$

From this we deduce that the charge conjugation operator is antiunitary and can be expressed as

$$\hat{\mathcal{C}} = U\hat{\mathcal{K}}_0 \tag{1.110}$$

where $\hat{\mathcal{K}}_0$ is the complex conjugation operator and U is some unitary operator defined by the relations

$$\begin{array}{rclcll} U\boldsymbol{\alpha}^*U^{-1} & = & U\left(\alpha_x, -\alpha_y, \alpha_z\right)U^{-1} & = & \boldsymbol{\alpha} & (a) \\ U\beta U^{-1} & = & -\beta & & & (b) \end{array} \tag{1.111}$$

Since the $\boldsymbol{\alpha}$-matrices transform as the coordinates, we see that the operator U is related to the operation of reflection in the xz-plane Eq.(1.66). However, all reflections commute with the β-matrix and to fulfill condition Eq.(1.111b) a slight modfication is required. With the introduction of an arbitrary phase we arrive at

$$\hat{\mathcal{C}} = i\beta\alpha_y\hat{\mathcal{K}}_0 \tag{1.112}$$

The effect of the charge conjugation operator on a 4-spinor is

$$\hat{\mathcal{C}}\psi(E) = i\beta\alpha_y\hat{\mathcal{K}}_0 \begin{bmatrix} \psi^{L\alpha} \\ \psi^{L\beta} \\ \psi^{S\alpha} \\ \psi^{S\beta} \end{bmatrix} = \psi^C(-E) \begin{bmatrix} \psi^{S\beta*} \\ -\psi^{S\alpha*} \\ -\psi^{L\beta*} \\ \psi^{L\alpha*} \end{bmatrix} \tag{1.113}$$

In particular, if ψ is a free-electron solution of energy +E, then ψ^C is a free-electron solution of energy −E.

Let us now look into the physics of charge conjugation. In the early days of quantum mechanics, the negative energy solutions was an extremely troublesome aspect of the Dirac equation. Contrary to classical mechanics, the negative energy solutions can not be discarded, since there will always be a finite probability for transitions between positive and negative energy solutions. However, this makes atoms and molecules highly unstable, contrary to what is observed experimentally. It can for instance be shown that the hydrogen atom would decay in less than one nanosecond and the electron would cause a radiative catastrophe as it falls down the negative energy branch through an infinite succession of transitions[84]. Dirac provided a solution [85] by stating that all negative energy states are occupied. Transitions down into the negative energy branch, the "Dirac sea", are then forbidden by the Pauli exclusion principle. On the other hand, transitions from the negative energy branch into positive energy states may be induced by energies larger than $2mc^2$:

$$E_{\text{positive energy electron}} - E_{\text{negative energy electron}} > 2mc^2 \tag{1.114}$$

However, from charge conjugation symmetry we see that we can reinterpret the negative energy electronic states as positive energy states of a particle of electron mass, but with charge +e, which will be termed the *positron*. It is the antiparticle of the electron, the "hole" left by the excited electron. Excitations of an electron out of the Dirac sea is then seen to be the creation of an electron-positron pair:

$$E_{\text{electron}} + E_{\text{positron}} > 2mc^2 \tag{1.115}$$

The positron was observed experimentally in 1932 [86]. This model leads to a radical reinterpretation of the vacuum. The energy-time uncertainity relation allows the creation of virtual electron-positron pairs at energies below $2mc^2$ so that the vacuum is now a "bubbling soup" of virtual pairs that polarize in the presence of external fields. Its proper mathematical description is provided by quantum electrodynamics (QED), which allows particle numbers to vary. In this theory, electrons and positrons appear as quanta of the quantized Dirac field, and vacuum fluctuations are zero-point oscillations of this field. Electromagnetic interactions are mediated by the exchange of virtual photons of the correspondingly quantized electromagnetic field. The interaction of the electron with zero-point oscillations of the electromagnetic field constitute its self energy. The combined effect of *vacuum polarization* and *self energy* is observed experimentally as the *Lamb shift* [87, 88], which in the hydrogen atom leads to a splitting of the $2\mathrm{s}_{1/2}$ and $2\mathrm{p}_{1/2}$ - atomic levels by $0.035\mathrm{cm}^{-1}$ (about 10% of the spin-orbit splitting of 2p). The Lamb shift is of order $\alpha(Z\alpha)^4$ where $\alpha \approx (1/137)$ is the fine-structure constant. The scaling to fourth order in the nuclear charge Z means that the Lamb shift can become quite large in high-Z systems. In U^{91+} the splitting due to the Lamb shift is thus of the order $6 \cdot 10^5\mathrm{cm}^{-1}$, as has been observed experimentally[89]. In relativistic molecular calculations we will invoke the *no-pair approximation*, that is we shall neglect all QED effects, which means that our theory can only be correct to the order of the Lamb shift.

1.7 The Dirac equation in a molecular field

> *... John Dalton ... used square wooden blocks to explain the atomic theory, with the result that a 'dunce, when asked to explain the atomic theory, said: "Atoms are square blocks of wood invented by Mr Dalton"'*
>
> T.TRAVIS & T.BENFEY *(1992) [90]*

After having explored the symmetries of the general Dirac equation in the previous three sections, we now construct a Dirac operator pertinent to the problem at hand, namely the description of molecules. For the moment we will restrict our attention to the one-electron part of the problem, and defer the discussion of the electron-electron interaction until the next chapter. Furthermore, we shall for the moment only consider stationary states of molecular systems. This means that the operator $i\frac{\partial}{\partial t}$ is replaced by the stationary energy E after extraction of a complex phase $\exp(-iEt)$. In the spirit of the Born-Oppenheimer approximation [91] the nuclei will be treated as stationary sources of external fields. This fixes our frame of reference and thereby ruins Lorentz invariance. It does, however, provide a workable model for the relativistic description of molecules. Let us therefore consider the electromagnetic fields generated by static nuclei.

If a nucleus posesses spin $\mathbf{I}$, it generates a vector potential

$$\mathbf{A}_n(\mathbf{r}) = \alpha^2 g_N \mu_N \frac{\mathbf{I}\times\mathbf{r}}{r^3} \tag{1.116}$$

where the nuclear g-factor g_N is of order unity and α is the fine-structure constant. The ratio between the nuclear magneton μ_N and the corresponding Bohr magneton μ_B for the electron is inversely proportional to the ratio of masses between the proton and the electron. Due to this ratio the magnetic fields generated by nuclei are about thousand times smaller than the fields generated by the electron spin magnetic moment. The magnetic *hyperfine* effects generated by nuclear spins may therefore safely be treated as perturbations. Magnetic fields from other external sources in a typical experimental situation are in general even smaller. We may therefore neglect the vector potential in variational calculations. This has the advantage that the resulting Dirac operator is time reversal symmetric and that its solutions are thereby at least doubly degenerate. As seen in section 1.5, this may be exploited in a quaternion formulation of the Dirac equation.

The scalar potential due to a nucleus N has the general form

$$\phi(\mathbf{r}_i) = \int \frac{\rho_N(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}' \tag{1.117}$$

where ρ_N is the nuclear charge distribution. In non-relativistic theory, nuclei are usually treated as point charges

$$\rho_N(\mathbf{r}_i) = Z_N \delta(\mathbf{r}_N) \tag{1.118}$$

(where δ is the three-dimensional Dirac delta function) giving a Coulombic potential

$$\phi(\mathbf{r}_i) = \frac{Z_N}{r_{iN}}; \qquad \mathbf{r}_{iN} = \mathbf{r}_i - \mathbf{r}_N \tag{1.119}$$

However, this introduces weak[10] singularities in the relativistic wave functions at the nuclear origins [92], which are hard to model in a finite basis approximation. With the introduction of a finite nucleus model, which in any case is a better physical model, the wave functions attain a Gaussian shape at the nuclear origins [93]. This favors the use of Gaussian basis functions in a finite basis approximation, which is the standard choice in non-relativistic molecular calculations. The nuclear charge distribution itself is often represented by a single Gaussian function [94]

$$\rho_N(\mathbf{r}_i) = Z_N \left(\frac{\eta_N}{\pi}\right)^{\frac{3}{2}} \exp\left[-\eta_N \mathbf{r}_{iN}^2\right] \tag{1.120}$$

since this leads to integrals that are easily evaluated. The exponent η_N is chosen to give a root-mean-square value $\langle r \rangle^{\frac{1}{2}}$ of the nuclear charge distribution

$$\left\langle r^2 \right\rangle^{\frac{1}{2}} = \frac{3}{2\eta} \tag{1.121}$$

equal to the empirical formula [95, 13]

$$\left\langle r^2 \right\rangle^{\frac{1}{2}} = \left[0.836 A^{\frac{1}{3}} + 0.57\right] \cdot 10^{-15} m \text{bohr} \tag{1.122}$$

where A is the atomic mass number. This gives the formula

$$\eta = \frac{3}{2} \left[\frac{0.529167}{0.836 A^{\frac{1}{3}} + 0.57}\right]^2 \cdot 10^{10} \tag{1.123}$$

The corresponding potential is then given in terms of the incomplete gamma functions F_n:

$$\phi(\mathbf{r}_i) = Z_N \sqrt{\frac{4\eta_N}{\pi}} F_0(\eta_N \mathbf{r}_{iN}^2); \qquad F_n(x) = \int_0^1 \exp\left[-x t^2\right] t^{2n} dt \tag{1.124}$$

The Gaussian charge distribution leads to a long-range behaviour of the potential identical to that of a Coulombic potential, but a finite value at the nuclear origin

$$\phi(\mathbf{r}_N) = 2 Z_N \sqrt{\frac{\eta_N}{\pi}} \tag{1.125}$$

---

[10] The singularities are weak in the sense that the wave function is still square integrable.

The shape of the nuclear charge distribution is not considered important for chemistry [96], but the Gaussian model can easily be improved by including more Gaussian functions. For instance, a nuclear quadrupole moment could be introduced by inclusion of Gaussian d-functions in order to model electric hyperfine interactions.

Based on the discussion above, we see that we can choose the time reversal symmetric Dirac operator to describe an electron in a molecular field. The molecular field Dirac operator can thereby be given a compact form in terms of quaternion algebra.

1.8 Approximations to the Dirac equation

> *The number 1/137, the fine structure constant which his teacher Sommerfeld had introduced into physics, was Pauli's link to the 'magic-symbolic' world with which he was so familiar. Pauli spent the last few days of his life in the Red Cross Hospital in Zürich, where he died on 15 December 1958. A fact which had disturbed him during these last days was that the number of his room was 137.*
>
> C.P.Enz*(1972) [97]*

1.8.1 4-component forms

In this section we consider the approach to the non-relativistic limit (n.r.l.) $c \rightarrow \infty$ of the Dirac equation for molecular fields discussed in the previous section. All changes in the values of observables that result from switching from the Dirac equation to its non-relativistic counterpart constitute *relativistic effects*. By a perturbation expansion of the Dirac equation in parameters related to the fine structure constant $\alpha = c^{-1}$, it is possible to obtain approximate Hamiltonians that incorporate relativistic effects to a given order in the expansion parameter. The purpose of this section is not to provide a comprehensive overview of such approximate Hamiltonians. Rather we employ perturbation expansion to identify and investigate the physical content of various relativistic effects. In addition, we shall discuss the difficulties associated with the derivation of approximate Hamiltonians.

Let us first establish the n.r.l. of the Dirac equation. We shall limit our discussion to the Dirac equation of an electron in the molecular field of fixed nuclei. To align the relativistic and non-relativistic energy scales, we subtract the rest mass term mc^2 from the Dirac equation. This amounts to the substitution

$$\beta \rightarrow \beta' = \beta - I_4 \tag{1.126}$$

and leads to the operator

$$\hat{h}_{D;V} = \left[\beta' mc^2 + c\,(\boldsymbol{\alpha}\cdot\mathbf{p}) + \hat{V}\right]; \qquad \hat{V} = -e\phi \tag{1.127}$$

and the corresponding equation

$$\begin{bmatrix} \hat{V} & 0 & -icd_z & -icd_- \\ 0 & \hat{V} & -icd_+ & icd_z \\ -icd_z & -icd_- & \hat{V}-2mc^2 & 0 \\ -icd_+ & icd_z & 0 & \hat{V}-2mc^2 \end{bmatrix} \begin{bmatrix} \psi^{L\alpha} \\ \psi^{L\beta} \\ \psi^{S\alpha} \\ \psi^{S\beta} \end{bmatrix} = E \begin{bmatrix} \psi^{L\alpha} \\ \psi^{L\beta} \\ \psi^{S\alpha} \\ \psi^{S\beta} \end{bmatrix} \tag{1.128}$$

Alternatively, it can be expressed as two coupled first-order differential equations

$$\begin{array}{rcrcll} \hat{V}\psi^L & + & c\,(\boldsymbol{\sigma}\cdot\mathbf{p})\,\psi^S & = & E\psi^L; & (a) \\ c\,(\boldsymbol{\sigma}\cdot\mathbf{p})\,\psi^L & + & \left(\hat{V}-2mc^2\right)\psi^S & = & E\psi^R; & (b) \end{array} \tag{1.129}$$

By solving Eq.(1.129b) for ψ^S, we find that the large ψ^L and small ψ^S components are coupled through the relation

$$\psi^S = \frac{1}{2mc}B(E)\,(\boldsymbol{\sigma}\cdot\mathbf{p})\,\psi^L; \qquad B(E) = \left[1+\frac{E-V}{2mc^2}\right]^{-1} \tag{1.130}$$

The energy-dependent operator $B(E)$ is totally symmetric under the symmetry group defined by the molecular field. It is therefore the operator $(\boldsymbol{\sigma}\cdot\mathbf{p})$ that relates the symmetries of the large and small components. The large and small components are thereby seen to have opposite parities, and we may anticipate that in a finite basis approximation to the Dirac equation, the large and small components will have to be expanded in separate basis sets.

Electronic solutions of the Dirac equation Eq.(1.129) have energies $E \sim 0$, which means that the small components are generally of order α smaller than the large components and vanish in the n.r.l., hence the notation employed for the upper and lower two components. The roles are reversed for positronic solutions ($E \sim -2mc^2$). Note, however, that for singular potentials such as the Coulombic potential the n.r.l. is only reached asymptotically and even for a finite nucleus Eq.(1.130) indicates that the small components of electronic solutions are to a large extent localized at the nuclear origins.

It is obvious from the form of Eq.(1.129) that its non-relativistic limit cannot be obtained directly. However, we may follow the approach of Kutzelnigg [98] and perform a change of metric by the substitution

$$\begin{bmatrix} \psi^L \\ \psi^S \end{bmatrix} \rightarrow \begin{bmatrix} \psi^L \\ \phi^L \end{bmatrix} = \begin{bmatrix} \psi^L \\ c\psi^S \end{bmatrix} \tag{1.131}$$

In matrix form we obtain the equation

$$\begin{bmatrix} \hat{V} & (\boldsymbol{\sigma}\cdot\mathbf{p}) \\ (\boldsymbol{\sigma}\cdot\mathbf{p}) & -2m\left[1-\dfrac{\hat{V}}{2mc^2}\right] \end{bmatrix} \begin{bmatrix} \psi^L \\ \phi^L \end{bmatrix} = E \begin{bmatrix} 1 & 0 \\ 0 & c^{-2} \end{bmatrix} \begin{bmatrix} \psi^L \\ \phi^L \end{bmatrix} \tag{1.132}$$

which for electronic solutions goes into the 4-component non-relativistic Lèvy-Leblond equation Eq.(1.42) in the non-relativistic limit. Note that in this approach the upper and lower two components are of the same order in α, which we have indicated by using the superscript L on both components. The n.r.l. for positronic solutions is obtained by the analogous substitution

$$\begin{bmatrix} \psi^L \\ \\ \psi^S \end{bmatrix} \rightarrow \begin{bmatrix} \phi^S \\ \\ \psi^S \end{bmatrix} = \begin{bmatrix} c\psi^L \\ \\ \psi^S \end{bmatrix} \tag{1.133}$$

The fact that positronic and electronic solutions have separate non-relativistic limits may be exploited to impose boundary conditions on approximate Hamiltonians, so that they are restricted to electronic solutions only. The modified Dirac equation Eq.(1.132) forms the basis for the direct perturbation theory of Kutzelnigg [98, 14].

The n.r.l. of the modified Dirac equation clearly demonstrates that spin is not a relativistic effect. The non-relativistic spin-free Hamiltonian of an electron in a molecular field is, however, straightforwardly obtained from the non-relativistic Lèvy-Leblond equation by elimination of the small components. In the non-relativistic domain, the spatial and spin degrees of freedom can therefore be treated separately. A separation of spin-free and spin-dependent terms is possible in the Dirac equation as well, as shown by Dyall [99]. His approach involves a change of metric first suggested by Kutzelnigg [100]

$$\begin{bmatrix} \psi^L \\ \\ \psi^S \end{bmatrix} \rightarrow \begin{bmatrix} \psi^L \\ \\ \phi^L \end{bmatrix} \qquad \text{where } \psi^S = \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})}{2mc}\phi^L \tag{1.134}$$

and leads to two coupled second-order differential equations

$$\begin{bmatrix} \hat{V} & \hat{T} \\ \\ \hat{T} & \left\{\dfrac{(\boldsymbol{\sigma}\cdot\mathbf{p})\,\hat{V}\,(\boldsymbol{\sigma}\cdot\mathbf{p})}{4m^2c^2} - \hat{T}\right\} \end{bmatrix} \begin{bmatrix} \psi^L \\ \\ \phi^L \end{bmatrix} = E \begin{bmatrix} 1 & 0 \\ \\ 0 & \dfrac{\hat{T}}{2mc^2} \end{bmatrix} \tag{1.135}$$

involving the kinetic energy operator

$$\hat{T} = \frac{p^2}{2m} \tag{1.136}$$

The Hamiltonian of this modified Dirac equation can be split into a spin-free $\tilde{h}^{sf}_{D;V}$ and a spin-dependent $\tilde{h}^{sd}_{D;V}$ part using Eq.(1.38):

$$(\boldsymbol{\sigma}\cdot\mathbf{p})\,\hat{V}\,(\boldsymbol{\sigma}\cdot\mathbf{p}) = (\boldsymbol{\sigma}\cdot\mathbf{p})\left(\boldsymbol{\sigma}\cdot\hat{V}\mathbf{p}\right) = \mathbf{p}\cdot\hat{V}\mathbf{p} - i\sigma\cdot\left(\mathbf{p}\times\hat{V}\mathbf{p}\right) \tag{1.137}$$

We obtain

$$\tilde{h}^{sf}_{D;V} = \begin{bmatrix} V & T \\ T & \dfrac{\mathbf{p}\hat{V}\cdot\mathbf{p}}{4m^2c^2} \end{bmatrix}; \quad \tilde{h}^{sd}_{D;V} = \begin{bmatrix} 0 & 0 \\ 0 & -\dfrac{i\sigma\cdot\left(\mathbf{p}\hat{V}\times\mathbf{p}\right)}{4m^2c^2} \end{bmatrix} \tag{1.138}$$

An advantage of this approach is that the functions ψ^L and ϕ^L have the same symmetry and may therefore be expanded in the same basis in a finite basis approximation.

1.8.2 Reduction to 2-component forms

The majority of approximate methods are based on a reduction of the Dirac equation to two-component form. This is accomplished either by elimination of the small components (esc) or by a unitary transformation that seeks to decouple the large and small components, e.g. the Foldy-Wouthuysen(FW) transformation[101]. It has been shown by Moss that the two approaches are equivalent for free particles, in which case the FW - transformation is available in closed form. We may straightforwardly generalize this conclusion to include external fields [102]. Our exposition and notation follows closely that of van Lenthe et al. [103].

Elimination of the small components

Consider first the method of elimination of the small components. The Dirac equation can be cast in the form

$$\hat{h}_D\Psi = E\Psi \quad \Rightarrow \quad \begin{bmatrix} \hat{h}_{11} & \hat{h}_{12} \\ \hat{h}_{21} & \hat{h}_{22} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = E \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \tag{1.139}$$

which may be written as a pair of coupled equations:

$$\begin{aligned} \hat{h}_{11}\phi_1 + \hat{h}_{12}\phi_2 &= E\phi_1 \quad (a) \\ \hat{h}_{21}\phi_1 + \hat{h}_{22}\phi_2 &= E\phi_2 \quad (b) \end{aligned} \tag{1.140}$$

We solve Eq.(1.140b) for ϕ_2

$$\phi_2 = \left(E - \hat{h}_{22}\right)^{-1}\hat{h}_{21}\phi_1 = \chi\phi_1 \tag{1.141}$$

and insert the result into Eq.(1.140a), so that we obtain

$$\hat{h}^{esc}\phi_1 = \left[\hat{h}_{11} + \hat{h}_{12}\left(E - \hat{h}_{22}\right)^{-1}\hat{h}_{21}\right]\phi_1 = \left[\hat{h}_{11} + \hat{h}_{12}\chi\right]\phi_1 = E\phi_1 \tag{1.142}$$

Note that $\hat{h}^{esc}$ is energy-dependent, but the equation may be solved iteratively. This does not, however, decouple our equations: After solving Eq.(1.142), ϕ_1 and E may be inserted in Eq.(1.141) to obtain ϕ_2. A decoupling may be accomplished by a *renormalization.* This step is furthermore important in order to ensure an effective Hamiltonian that is Hermitian. Let Ψ be normalized to unity. We seek a normalized function of the form $\Phi = \hat{O}\phi_1$ where $\hat{O}$ is a normalization operator:

$$\langle \Phi \mid \Phi \rangle = \left\langle \hat{O}\phi_1 \mid \hat{O}\phi_1 \right\rangle = \langle \Psi \mid \Psi \rangle = \langle \phi_1 \mid \phi_1 \rangle + \langle \phi_2 \mid \phi_2 \rangle = 1 \tag{1.143}$$

Using Eq.(1.141) the final term is recast as

$$\langle \phi_1 \mid \phi_1 \rangle + \langle \phi_2 \mid \phi_2 \rangle = \left\langle \phi_1 \left| (1 + \chi^\dagger \chi) \right| \phi_1 \right\rangle \tag{1.144}$$

We may therefore choose $\hat{O}$ as

$$\hat{O} = \sqrt{1 + \chi^\dagger \chi} \tag{1.145}$$

and the Hamiltonian for Φ becomes

$$\hat{h}^{eff} = \hat{O}\hat{h}^{esc}\hat{O}^{-1} = \sqrt{1 + \chi^\dagger \chi}\left[\hat{h}_{11} + \hat{h}_{12}\chi\right]\frac{1}{\sqrt{1 + \chi^\dagger \chi}} \tag{1.146}$$

This, then, is the effective two-component Hamiltonian in its final form.

Decoupling by unitary transformation

The above decoupling may be accomplished through a unitary transformation as well.

$$U\hat{h}_D U^{-1} U \Phi = E U \Phi \tag{1.147}$$

The unitary operator U may be expressed as[14]

$$U = \hat{O}^{-1}\left[\begin{array}{cc} 1 & \chi^\dagger \\ -\chi & 1 \end{array}\right] \quad \Rightarrow \quad U^{-1} = \hat{O}^{-1}\left[\begin{array}{cc} 1 & -\chi^\dagger \\ \chi & 1 \end{array}\right]; \qquad \hat{O} = \sqrt{1 + \chi^\dagger \chi} \tag{1.148}$$

Our transformed Hamiltonian becomes

$$\begin{aligned} U\hat{h}_D U^{-1} &= \left[\begin{array}{ll} \hat{h}'_{11} & 0 \\ 0 & \hat{h}'_{22} \end{array}\right] \\ &= \hat{O}^{-1}\left[\begin{array}{ll} \hat{h}_{11} + \hat{h}_{12}\chi + \chi^\dagger \hat{h}_{21} + \chi^\dagger \hat{h}_{22}\chi & -\hat{h}_{11}\chi - \chi^\dagger \hat{h}_{21}\chi^\dagger + \hat{h}_{21} + \chi^\dagger \hat{h}_{22} \\ -\chi \hat{h}_{11} - \chi \hat{h}_{21}\chi + \hat{h}_{21} + \hat{h}_{22}\chi & \chi \hat{h}_{11}\chi^\dagger - \chi \hat{h}_{12} - \hat{h}_{21}\chi^\dagger + \hat{h}_{22} \end{array}\right]\hat{O}^{-1} \end{aligned} \tag{1.149}$$

From this we see that the conditions for decoupling of the two equations are

$$\begin{aligned} -\chi\hat{h}_{11} - \chi\hat{h}_{12}\chi + \hat{h}_{21} + \hat{h}_{22}\chi &= 0 \qquad \text{(a)} \\ -\hat{h}_{11}\chi^\dagger + \hat{h}_{12} - \chi^\dagger\hat{h}_{21}\chi^\dagger + \chi^\dagger\hat{h}_{22} &= 0 \qquad \text{(b)} \end{aligned} \tag{1.150}$$

The two equations are seen to be simply Hermitian conjugates of each others. By multiplying Eq.(1.150) equation with $\chi^\dagger$ from the left and rearranging we obtain

$$\chi^\dagger\hat{h}_{21} + \chi^\dagger\hat{h}_{22}\chi = \chi^\dagger\chi\hat{h}_{11} + \chi^\dagger\chi\hat{h}_{12}\chi \tag{1.151}$$

The latter relation may be used to simplify the upper diagonal transformed Hamiltonian

$$\begin{aligned} \hat{h}'_{11} &= \tfrac{1}{\sqrt{1+\chi^\dagger\chi}}\left[\hat{h}_{11} + \hat{h}_{12}\chi + \chi^\dagger\hat{h}_{21} + \chi^\dagger\hat{h}_{22}\chi\right]\tfrac{1}{\sqrt{1+\chi^\dagger\chi}} \\ &= \sqrt{1+\chi^\dagger\chi}\left[\hat{h}_{11} + \hat{h}_{12}\chi\right]\tfrac{1}{\sqrt{1+\chi^\dagger\chi}} \end{aligned} \tag{1.152}$$

We see that this is exactly the same expression as Eq.(1.146) if we choose

$$\chi = \left(E - \hat{h}_{22}\right)^{-1}\hat{h}_{21} \tag{1.153}$$

Thus the two decoupling schemes are equivalent.

1.8.3 2-component forms

The equivalence of the two decoupling schemes means that when one considers approximations one can choose the scheme that is best suited to the mathematical manipulations involved. The expression for the effective two-component Hamiltonian Eq.(1.146) is somewhat deceptive in that it gives the impression that the large and small components can be completely separated.

This is, however, only true for free particles and not in the presence of an external field. The effective two-component Hamiltonian derived in the previous section has the form

$$\hat{h}^{eff} = \hat{O}\hat{h}^{esc}\hat{O}^{-1} = \sqrt{1+\chi^\dagger\chi}\left[\hat{V} + \frac{1}{2m}\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)B(E)\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\right]\frac{1}{\sqrt{1+\chi^\dagger\chi}} \tag{1.154}$$

where

$$\chi = \frac{1}{2mc}B(E)\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right) \tag{1.155}$$

By a slight rearrangement we obtain

$$\hat{h}^{eff} = \hat{h}^{esc} - \left[\hat{h}^{esc}, \hat{O}\right] \hat{O}^{-1} \tag{1.156}$$

In the case of free particles the normalization operator $\hat{O}$ commutes with $\hat{h}^{esc}$, so that the second term in the effective Hamiltonian Eq.(1.156) disappears. The remaining equation reads

$$\hat{h}^{eff}\Phi = \hat{h}^{esc}\Phi = c^2\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\left(E + 2mc^2\right)^{-1}\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\Phi = \frac{p^2}{2m}\left[1 + \frac{E}{2mc^2}\right]^{-1}\Phi = E\Phi \tag{1.157}$$

We rearrange to

$$\left(E^2 + 2mc^2E - c^2p^2\right)\Phi = 0 \tag{1.158}$$

and solve for E, which gives

$$\hat{h}^{eff}_{\pm}\Phi = \left(\pm\hat{E}_p - mc^2\right)\Phi_{\pm} = E_{\pm}\Phi_{\pm}; \qquad \hat{E}_p = \sqrt{m^2c^4 + c^2p^2} \tag{1.159}$$

The sign refers to free electron (+) or positron (−) solutions. The square root in the operator represents no problem, since we can solve this equation in momentum space.

With the introduction of external fields the normalization operator $\hat{O}$ does not in general commute with $\hat{h}^{esc}$ and the effective two-component Hamiltonian can not be expressed in an operationally useful form. It has to be approximated by some sort of perturbation expansion from which the energy dependence can be iterated out. The major difficulty with such an approach is that it tends to give operators that are ill-behaved. A simple example is provided by the effective two-component Hamiltonian for free electrons. An expansion to second order in $(p/mc)^2$ gives

$$\hat{h}^{eff}_{+} = mc^2\left(\sqrt{1 + \left(\frac{p}{mc}\right)^2} - 1\right) = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + O\left(\alpha^4\right) \tag{1.160}$$

The first term in the expansion is the standard kinetic energy operator $\hat{T}$. By comparison with Eq.(1.28) we recognize the second term as a first order correction to the kinetic energy resulting from the increase of the electron mass with the velocity. It is therefore commonly referred to as the *mass-velocity* term. There are two difficulties associated with this expansion:

1. The expansion in $(p/mc)^2$ is valid only for $p \ll mc$ which in general does not hold true. In particular, $p \to \infty$ as $r \to 0$ in the presence of a Coulombic potential.

2. Even though $\hat{h}_+^{eff}$ is defined for free electrons only, i.e. for energies $E > 0$, the operator resulting from the second order expansion above is unbounded from below due to the presence of the mass-velocity term. The inclusion of higher order terms lead to uncontrollable oscillations.

Let us briefly look at three effective Hamiltonians for an electron in an external molecular field:

1. The Pauli operator.
2. The regular approximation.
3. The Douglas-Kroll operator.

The three effective Hamiltonians can be related to three different expansions of the energy-dependent operator

$$B(E) = \left[1 + \frac{E - V}{2mc^2}\right]^{-1} \tag{1.161}$$

The Pauli-operator

In the first approach $B(E)$ is expanded using the series expansion

$$(1 + x_1)^{-1} = \sum_{k=0}^{\infty} (-1)^k \, x_1^k; \qquad x_1 = \frac{E - V}{2mc^2} \tag{1.162}$$

The Pauli-operator [52] is obtained by an expansion to second order in x_1 and has the form

$$\hat{h}^P = \hat{h}_{nr} + \hat{h}_{mv} + \hat{h}_{so} + \hat{h}_{Dar} \tag{1.163}$$

The first term is simply the non-relativistic Hamiltonian

$$\hat{h}_{nr} = \hat{T} + \hat{V} \tag{1.164}$$

whereas the second term is the mass-velocity operator

$$\hat{h}_{mv} = -\frac{p^4}{8m^3c^2} \tag{1.165}$$

The third term is the *spin-orbit operator* [104, 105, 106, 16]

$$\hat{h}_{so} = \frac{ie\boldsymbol{\sigma}}{4m^2c^2} \cdot (\mathbf{E} \times \boldsymbol{\nabla}) \tag{1.166}$$

where $\mathbf{E}$ is the molecular field. Its physical content can be understood by comparison with Eq.(1.26). When an electron moves in the molecular field it experiences a magnetic field in its own rest frame, in addition to the electric field. The spin-orbit operator represents the interaction of the electron spin with this magnetic field, as mediated by the electronic motion. The operator couples spin and spatial degrees of freedom and therefore makes a separate treatment of spin and spatial symmetry impossible.

The fourth term is the *Darwin term*

$$-\frac{e}{8m^2c^2}\left(\boldsymbol{\nabla}\cdot\mathbf{E}\right) \tag{1.167}$$

and can be understood as a correction to the electric potential of the nuclei felt by the electron if we assume that the electron (or its charge) performs a rapid isotropic oscillatory motion $\boldsymbol{\delta}$ (*Zitterbewegung*) [107] about its mean position $\mathbf{r}$. In order to see this, we expand the potential about the mean position $\mathbf{r}$

$$\phi(\mathbf{r}+\boldsymbol{\delta}) = \phi(\mathbf{r}) + (\boldsymbol{\delta}\cdot\boldsymbol{\nabla})\phi(\mathbf{r}) + \frac{1}{2}(\boldsymbol{\delta}\cdot\boldsymbol{\nabla})^2\phi(\mathbf{r}) + \ldots \tag{1.168}$$

and take the time average[11] [40]

$$\begin{aligned} \langle\phi(\mathbf{r}+\boldsymbol{\delta})\rangle &= \phi(\mathbf{r}) + \tfrac{1}{2}\left\langle(\boldsymbol{\delta}\cdot\boldsymbol{\nabla})^2\right\rangle\phi(\mathbf{r}) + \ldots \\ &= \phi(\mathbf{r}) + \tfrac{1}{6}\left\langle\delta^2\right\rangle\nabla^2\phi(\mathbf{r}) + \ldots \end{aligned} \tag{1.169}$$

The Darwin term arises from the second term if we make the identification

$$\left\langle\delta^2\right\rangle = \frac{3}{(2mc)^2} \tag{1.170}$$

The extraordinary dynamics of the Dirac electron is illustrated by the fact that it has three sets of independent dynamical variables: position $\mathbf{r}$, momentum $\mathbf{p}$, and velocity $c\boldsymbol{\alpha}$. The latter operator is found from the Heisenberg equation of motion

$$\frac{d\mathbf{r}}{dt} = i\left[\hat{h}_D, \mathbf{r}\right] = c\boldsymbol{\alpha} \tag{1.171}$$

The extra degrees of freedom accomodates spin, and it has therefore been suggested that spin arises from the *Zitterbewegung*, interpreted as internal charge oscillations in the electron [108, 109, 110, 111].

The Pauli-Hamiltonian is not a very satisfactorily approximation to the Dirac equation for two reasons:

---

[11] The first order time in the Taylor expansion the disappears due to the assumed isotropy of the *Zitterbewegung*.

1. It contains the mass-velocity operator, which makes it unbounded from below and therefore not suited for variational calculations.

2. If we make the association $E - \hat{V} = \hat{T}$, we see that the expansion parameter x_1 is approximately $(p/mc)^2$ which goes to infinity as $r \to 0$ in singular potentials and therefore leads to the introduction of highly singular operators. In particular, the spin-orbit operator in a Coulombic potential $\hat{V} = -Z/r$ attains the form [12]

   $$\hat{h}_{so} = \frac{Z}{4m^2c^2}\left(\frac{\boldsymbol{\sigma}\cdot\mathbf{l}}{r^3}\right), \tag{1.172}$$

   which has an r^{-3} dependence. The Darwin term becomes a Dirac delta function

   $$\hat{h}_{Dar} = \frac{\pi}{2m^2c^2}\delta(\mathbf{r}) \tag{1.173}$$

   and thereby only contributes a positive energy shift determined by the electron density at the nuclear origins, which is not very satisfactorily in a variational calculation. Higher expansions in x_1 Eq.(1.162) leads to higher singularities and undefined products of three-dimensional delta functions. Note that the situation is not alleviated with the introduction of finite nuclei, for even though the expansion parameter no longer goes to infinity as $r \to 0$, it will still have very large values near the nuclear origins. This will be illustrated by a numerical example below.

Further discussion of the Pauli-operator is found in Refs. [113, 114, 4].

The regular approximation

Singularities in the effective two-component Hamiltonian can be avoided by a better choice of expansion parameter. A regular expansion has been suggested by van Lenthe et al. [103]

$$B(E) = \frac{2mc^2}{2mc^2 - V}\left(1 + x_2\right)^{-1}; \qquad x_2 = \frac{E}{2mc^2 - V} \tag{1.174}$$

The extraction of a prefactor from $B(E)$ means that the zero–order Hamiltonian in the regular approximation (ZORA)

$$\hat{h}^{ZORA} = \hat{V} + (\boldsymbol{\sigma}\cdot\mathbf{p})\left\{\frac{c^2}{2mc^2 - V}\right\}(\boldsymbol{\sigma}\cdot\mathbf{p}) \tag{1.175}$$

---

[12] Expressions for the spin-orbit and Darwin terms with a Gaussian nuclear charge distribution has been derived by Dyall and Fægri [112].

is not identical to the non-relativistic Hamiltonian. Rather, it is identical to the zero–order effective Hamiltonian derived by Chang et al. [115]. It can be reformulated as

$$\hat{h}^{ZORA} = \hat{V} + \mathbf{p}\left\{\frac{c^2}{2mc^2 - V}\right\}\mathbf{p} - \left\{\frac{c^2}{2mc^2 - V}\right\}^2 ie\boldsymbol{\sigma}\cdot(\mathbf{E}\times\boldsymbol{\nabla}) \tag{1.176}$$

By comparison with Eq.(1.166) this shows that spin-orbit interaction is present already at the zero–order level.

We may illustrate the difference between the two expansions Eq.(1.162) and Eq.(1.174) by a numerical example. For an atomic stationary state the $B(E)$ operator may be considered as an analytic function of the radial variable r [116]. This allows a direct comparison of expansions. I have chosen the one-electron system Rn^{85+} as test example. The nuclear charge distribution has been modelled by a Gaussian function Eq.(1.120) with exponent $\eta = 1.3\cdot 10^8$, corresponding to a root-mean-square value $\langle r\rangle = 1.1\cdot 10^{-4}$ bohr of the nuclear charge distributions. Using the atomic 4-component code GRASP [117], the energy of the $1s_{1/2}$ orbital was found to be $E_{1s_{1/2}} = -4154.662042H$[13].

Due to the use of a finite nucleus the potential does not go to infinity as $r \rightarrow 0$, but it does reach a large value

$$V(0) = -Z \times 2\sqrt{\frac{\eta}{\pi}} = -1.117\cdot 10^6 \tag{1.177}$$

This means that x_1 is not a very good expansion parameter near the nuclear origin. This is clearly displayed in Table 1.2 where I have listed the values of the potential, of $B(E)$ and the two expansion parameters x_1 and x_2 at the nuclear origin and at infinity. In Fig.1.1 the expansion parameters are plotted as functions of r. It can be seen that whereas x_2 goes more or less to zero at the nuclear origin, x_2 goes to a value that is about thirty times unity. Expansions up to order two are plotted in Figures 1.2, 1.3 and 1.4. It can be seen that the expansion in x_2 is essentially converged already to second order. Limiting values for the two expansions up to order five are given in Tables 1.3 and 1.4. As a measure of the general convergence I have calculated the integral [14]

$$\tau(x,n) = \int_0^\infty \frac{|B(r;E) - b(r;E,x,n)|}{r}dr \tag{1.178}$$

where $b(r;E,x,n)$ represents an expansion of $B(E)$ as a function of r in terms of expansion parameter x up to order n.

The test example clearly demonstrates the superiority of x_2 over x_1 as expansion parameter. The regular approximation generates operators that are never more singular

---

[13] This may be compared with the value $E_{1s_{1/2}} = -4158.424082H$ obtained with a point nucleus

[14] The integrals were evaluated using the numerical integration routine of GRASP [117].

than the external potential. However, the mass-velocity operator appears in the first order Hamiltonian, which means that it is has no lower bound and can not be used in variational calculations.

Table 1.2: Limiting values of the potential, expansion parameters and $B(E)$

	r = 0	$r \to \infty$
$\hat{V}$	$-1.117 \cdot 10^{6}$	0.000
B(E)	$3.266 \cdot 10^{-2}$	1.124
x_1	29.617	$-1.106 \cdot 10^{-1}$
x_2	$-3.600 \cdot 10^{-3}$	$-1.106 \cdot 10^{-1}$

Table 1.3: Expansion in x_1

Order	r = 0	$r \to \infty$	τ
0	1.0000	1.000	$2.400 \cdot 10^{1}$
1	$-2.862 \cdot 10^{1}$	1.111	$1.260 \cdot 10^{2}$
2	$8.485 \cdot 10^{2}$	1.123	$3.208 \cdot 10^{3}$
3	$-2.513 \cdot 10^{4}$	1.124	$8.839 \cdot 10^{4}$
4	$7.443 \cdot 10^{5}$	1.124	$2.492 \cdot 10^{6}$
5	$-2.204 \cdot 10^{7}$	1.124	$7.105 \cdot 10^{7}$

Table 1.4: Expansion in x_2

Order	r = 0	$r \to \infty$	τ
0	$3.254 \cdot 10^{-2}$	1.000	$1.781 \cdot 10^{1}$
1	$3.266 \cdot 10^{-2}$	1.110	1.945
2	$3.266 \cdot 10^{-2}$	1.123	$2.147 \cdot 10^{-1}$
3	$3.266 \cdot 10^{-2}$	1.124	$2.370 \cdot 10^{-2}$
4	$3.266 \cdot 10^{-2}$	1.124	$2.619 \cdot 10^{-3}$
5	$3.266 \cdot 10^{-2}$	1.124	$2.893 \cdot 10^{-4}$

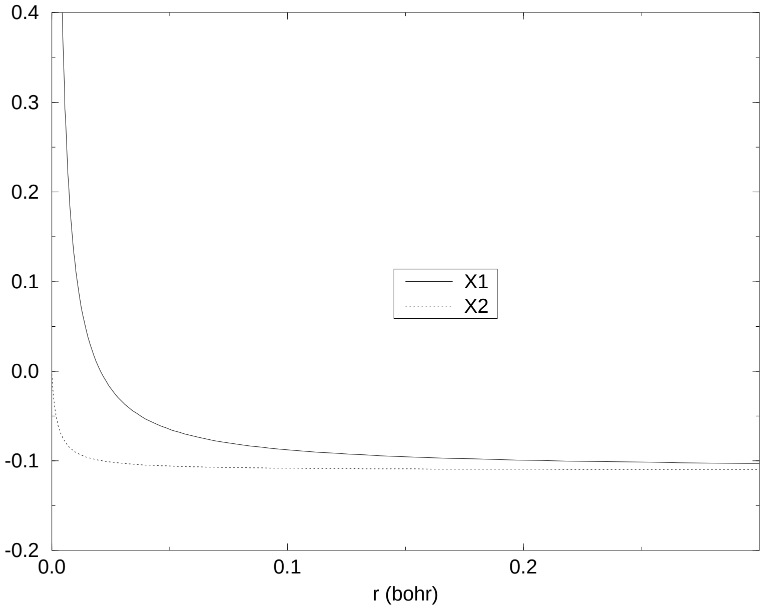


Figure 1.1: The expansion parameters x_1 Eq.(1.162) and x_2 Eq.(1.174) plotted as a function of r for the $1s_{1/2}$-orbital of Rn^{85+}.

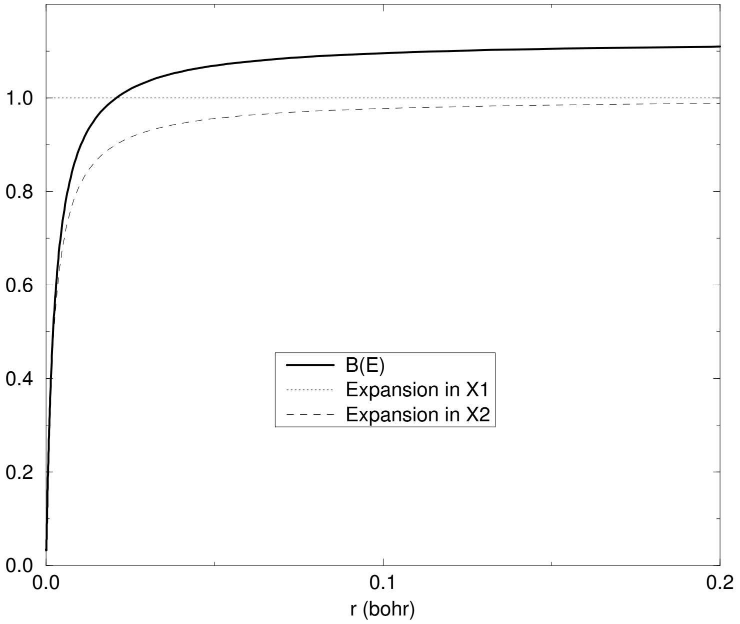


Figure 1.2: Comparison of the exact function $B(E)$ and zero–order expansions in x_1 Eq.(1.162) and x_2 Eq.(1.174) for the $1s_{1/2}$-orbital of Rn^{85+}.

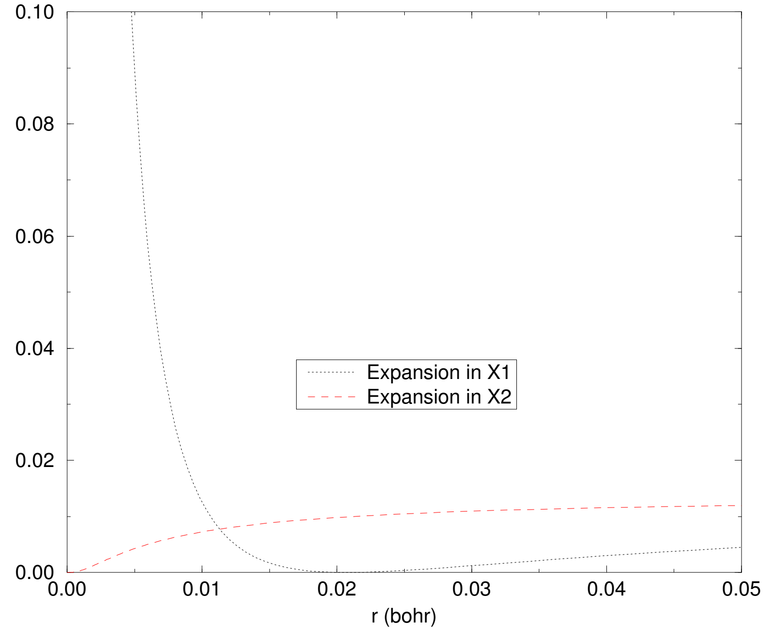


Figure 1.3: Deviations of first order expansions in x_1 Eq.(1.162) and x_2 Eq.(1.174) from the exact $B(E)$ for the $1s_{1/2}$-orbital of Rn^{85+}.

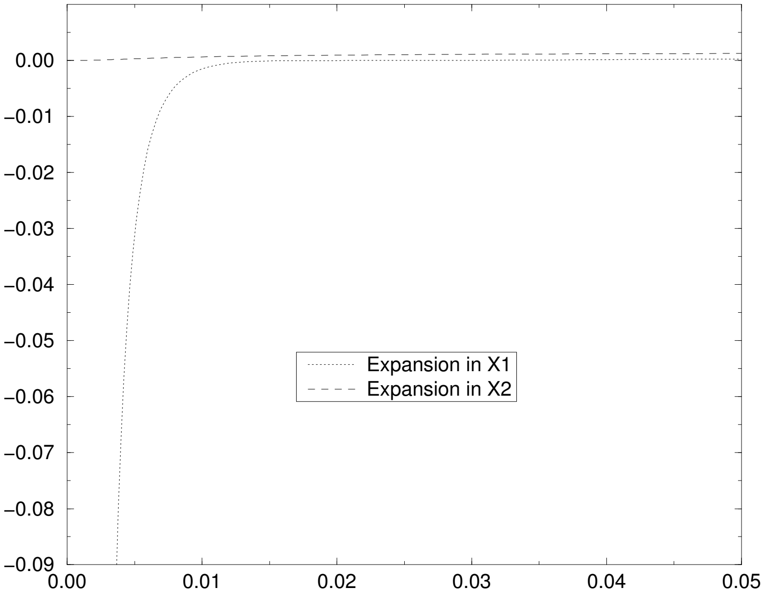


Figure 1.4: Deviations of second order expansions in x_1 Eq. (1.162)) and x_2 Eq. (1.174) from the exact $B(E)$ for the $1s_{1/2}$-orbital of Rn^{85+}.

The second order Douglas-Kroll operator

In order to avoid unboundedness of approximate Hamiltonians, we realize that the free-electron operator $\hat{E}_p$ Eq.(1.159) must be retained in closed form. Since this operator involves the square root of the momentum operator, the desired two-component Hamiltonian is therefore preferably derived in the momentum representation. The Douglas-Kroll operator [118] may be thought of as an expansion in

$$x_3 = \frac{E - V}{mc^2 + \hat{E}_p} \tag{1.179}$$

We see that $\hat{E}_p$ appears explicitly in the expansion parameter and is thereby never expanded. Furthermore, x_3 leads to a regular expansion of $B(E)$ since, in the presence of a Coulombic potential, $p \rightarrow \infty$ as $r \rightarrow 0$. The expansion therefore introduces no singularities more severe than the potential. Note that to first order in α^2 x_3 reduce to x_1, but the two expansion parameters cannot be compared directly since x_3 contains the non-local operator $\hat{E}_p$.

The Douglas-Kroll operator is formally derived by first performing a free-electron Foldy-Wouthuysen transformation on the Dirac operator and then expanding the resulting operator in the external potential.The operator has the form

$$\hat{h}^{DK} = \hat{E}_p - mc^2 + \hat{V}^{eff} \tag{1.180}$$

where the effective potential to second order is given by

$$\hat{V}^{eff} = \hat{A}\left[\hat{V} + \hat{R}\hat{V}\hat{R}\right]\hat{A} + \hat{W}_1\left[\hat{E}_p + \frac{1}{2}\left(\hat{W}_1\hat{E}_p + \hat{E}_p\hat{W}_1\right)\right]\hat{W}_1 \tag{1.181}$$

Kinematical factors are

$$\hat{A} = \sqrt{\frac{\hat{E}_p + mc^2}{2\hat{E}_p}} \tag{1.182}$$

introduced by renormalization, and

$$\hat{R} = \frac{c\,(\boldsymbol{\sigma}\cdot\mathbf{p})}{\hat{E}_p + mc^2} \tag{1.183}$$

which regularizes potentials since [119]

$$\lim_{p\rightarrow\infty}\hat{R} = (\boldsymbol{\sigma}\cdot\mathbf{n})\,; \qquad \mathbf{n} = \frac{\mathbf{p}}{|\mathbf{p}|} \tag{1.184}$$

Finally, $\hat{W}_1$ is an integral operator with kernel

$$\hat{W}_1\left(\mathbf{p},\mathbf{p}'\right) = \hat{A}(\mathbf{p})\left[\hat{R}(\mathbf{p}) - \hat{R}(\mathbf{p}')\right]\hat{A}(\mathbf{p}')\left\{\frac{\hat{V}(\mathbf{p},\mathbf{p}')}{E_p + E_{p'}}\right\} \tag{1.185}$$

Here $\hat{V}(\mathbf{p},\mathbf{p}')$ denotes the Fourier transform of the external potential. The Douglas-Kroll operator is not straightforwardly implemented due to use of the momentum representation. Applications to chemical systems have been made feasible through the work of Heß et al. [120, 15, 121].

Chapter 2

Many-electron systems

> *. . . there is an alternative formulation of the many-body problem, i.e. how many bodies are required before we have a problem ? G. E. Brown points out that this can be answered by a look at history. In eighteenth-century Newtonian mechanics, the three-body problem was insoluble. With the birth of general relativity around 1910 and quantum electrodynamics in 1930, the two- and one-body problems became insoluble. And within modern quantum field theory, the problem of zero bodies (vacuum) is insoluble. So, if we are out after exact solutions, no bodies at all is already too many !*
>
> R. D. MATTUCK*[122]*

A many-particle system represents no many-particle problem unless there is some sort of interaction between the particles. For an electron in a molcular field, we have so far avoided many-body problems by relegating nuclei to the role of stationary sources of external fields. With the introduction of electron-electron interactions we have a true many-body problem at hand, and as such it cannot in general be solved analytically[1]. In the relativistic domain a further complication is that the Lorentz invariant electron-electron interaction cannot be expressed in a useful closed form, which means that approximations must be introduced.

In this chapter we will first discuss the nature of the electron-electron interaction in relativistic systems. We then construct the Hamiltonian that will be used in relativistic calculations. To avoid overlap of material in the thesis I shall summarize the papers at this stage. The reader is adviced to go through the papers at this point (section 2.2.3). Before concluding, I will discuss various features of relativistic molecular calculations.

---

[1] An exactly soluble non-relativistic two-electron atomic model is discussed in [123].

2.1 The electron-electron interaction

> *A friend of ours, Hulme, was walking many years ago with Paul [Dirac] in Cambridge, with something rattling in his pocket, and he said apologetically, 'I am sorry about this noise. I have a bottle of aspirin in my pocket, and I took some as I had a cold, so it is normal for it to make some noise'. There was a silence for a while and then, so our friend Hulme told us, Paul said, 'I suppose it makes a maximum noise when it's just half full.'*
>
> R.Peierls *(1972) [61]*

In classical electrodynamics the interaction of two electrons is mediated by the electromagnetic field. In contrast, quantum electrodynamics (QED) views this interaction as mediated by the exchange of virtual photons. The link between the two pictures is that photons appear as quanta of the quantized electromagnetic field. The electrons themselves, together with their antiparticles (positrons), are quanta of the quantized Dirac or electron-positron field. An important consequence of field quantization is that it allows the number of quanta or particles to change, as for example in the creation or annihilation of electron-positron pairs. Relativistic molecular quantum mechanics in general operate within the framework of the no-pair approximation, in which pair creation is prohibited. This is equivalent to working with an unquantized Dirac field. Furthermore, the electron-electron interaction is usually represented solely by the instantaneous Coulomb interaction, which means that all retardation and direct magnetic effects are ignored. In this section we discuss the full electron-electron interaction of QED and to what extent the Coulomb interaction is a good approximation for our purposes. It will be seen that a change of gauge alters the form of the expression for electron-electron interaction in QED already at order $(Z\alpha)^2$. In order to understand how this comes about, it is instructive to explore the link between classical and quantum electrodynamics.

In classical electrodynamics the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields are determined by Maxwell's equations

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\mathbf{B} &= 0 && (a)\\
\boldsymbol{\nabla}\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t} && (b)\\
\boldsymbol{\nabla}\cdot\mathbf{E} &= 4\pi\rho && (c)\\
\boldsymbol{\nabla}\times\mathbf{B} &= \frac{1}{c^2}\left[4\pi\mathbf{j}+\frac{\partial\mathbf{E}}{\partial t}\right] && (d)
\end{aligned}\qquad(2.1)$$

where ρ is the charge density (charge per unit volume) and $\mathbf{j}$ is the current density (flow of charge per unit time per unit area across a surface). Maxwell's equations are Lorentz invariant and were therefore, prior to the advent of the theory of special relativity, denoted non-relativistic, since they did not obey the Galilean transformation Eq.(1.1). The electric and magnetic fields may be represented by a scalar ϕ and vector $\mathbf{A}$ potential

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}; \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{2.2}$$

Whereas the electric and magnetic fields are uniquely determined by the scalar and vector potentials, the inverse relation does not hold true. As mentioned in section 1.1, the electric and magnetic fields are invariant under gauge transformations

$$\mathbf{A} \rightarrow \mathbf{A} - \boldsymbol{\nabla} f; \qquad \phi \rightarrow \phi + \frac{\partial f}{\partial t} \tag{2.3}$$

where f is any scalar function of space and time coordinates. Specific choices of gauge are the *Coulomb gauge*

$$\boldsymbol{\nabla} \cdot \mathbf{A} = 0 \tag{2.4}$$

and the *Feynman (Lorentz) gauge*

$$\delta_\mu A_\mu = 0 \tag{2.5}$$

The Feynman gauge is expressed solely in terms of 4-vectors and is therefore manifestly Lorentz invariant.

The origin of the name "Coulomb gauge" is seen from the following: If we insert the definitions Eq.(2.2) of the electric and magnetic fields into Eq.(2.1c) and use the Coulomb gauge condition Eq.(2.4), we obtain

$$\boldsymbol{\nabla} \cdot \mathbf{E} = -\nabla^2\phi - \frac{\partial(\boldsymbol{\nabla} \cdot \mathbf{A})}{\partial t} = -\nabla^2\phi = 4\pi\rho \tag{2.6}$$

which is simply the Poisson equation with solution

$$\phi(\mathbf{r}, t) = \int \frac{\rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \tag{2.7}$$

The scalar potential in Coulomb gauge is thereby seen to be the instantaneous Coulomb potential due to a charge density $\rho(\mathbf{r}, t)$.

In the absence of charge and currents ($\rho = 0, \mathbf{j} = 0$), the scalar potential is identically zero and insertion of Eq.(2.2) in Eq.(2.1d) gives the equation

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = 0 \tag{2.8}$$

which describes the propagation of electromagnetic waves in free space. Solutions have the form

$$\mathbf{A}\left(\mathbf{r},t\right)=\mathbf{A}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-2\pi\nu t)} \tag{2.9}$$

where ν is the frequency. Direct application of the Coulomb gauge condition Eq.(2.4) gives

$$\mathbf{k}\cdot\mathbf{A}=0 \tag{2.10}$$

which means that the vector potential $\mathbf{A}$ Eq.(2.9) is perpendicular to the direction of propagation $\mathbf{k}$ of the wave. The vector potential is then said to be transversal. In Coulomb gauge the electric field splits up into a *transversal* $\mathbf{E}_\perp$ and *longitudinal* $\mathbf{E}_\|$ part defined by

$$\begin{array}{rclcrcl} \boldsymbol{\nabla}\cdot\mathbf{E}_\perp & = & 0 & \Rightarrow & \mathbf{E}_\perp & = & -\dfrac{\partial\mathbf{A}}{\partial t} \\ \boldsymbol{\nabla}\times\mathbf{E}_\| & = & 0 & \Rightarrow & \mathbf{E}_\| & = & -\nabla\phi \end{array} \tag{2.11}$$

The magnetic field $\mathbf{B}$ can only have a longitudinal contribution, which is explicitly shown by Eq.(2.1a). If we now insert the definition of the longitudinal electric field $\mathbf{E}_\|$ into Eq.(2.1c), we see, analogous to Eq.(2.6), that the longitudinal field is associated with the instantaneous Coulomb interaction and is uniquely determined by the charge distribution ρ. This also demonstrates that in Coulomb gauge all retardation and magnetic interactions enter the transverse part of the electromagnetic field. A disadvantage of the Coulomb gauge is that it is not Lorentz invariant, so that if one changes the frame of reference a gauge transformation is in general needed in order to reestablish the gauge conidtion Eq.(2.4).

The link between classical and quantum electrodynamics can be summed out by a short historic survey: At the end of the last century it was realized that the electromagnetic field could be treated as a collection of independent harmonic oscillators, each associated with a particular frequency ν. A major step towards quantum mechanics was the postulate made by Planck in 1900 [124] stating that the energy absorbed or emitted by such radiation oscillators were not continuous, but appeared in quanta of hν. This postulate was necessary in order to avoid the radiation catastrophe in blackbody radiation. In 1905 Einstein explained the photoelectric effect by assuming that the electromagnetic field could be considered as a collection of independent energy quanta[2] of magnitude $h\nu$. Then, in 1927, Dirac laid the foundations of quantum electrodynamics by introducing creation and annihilation operators of photons to describe the interaction of light with matter. Finally, in 1932 Bethe and Fermi described electromagnetic interaction in terms of virtual photon exchange. This lead to expressions for the electron-electron interaction in

---

[2] The word 'photon' was first introduced in 1926 by G.N.Lewis [125].

configuration space which depends on the momentum ω transferred by the exchanged photon [126, 127, 11, 128, 129, 130]. In the Coulomb gauge the electron-electron interaction $\hat{g}_\omega(1,2)$ (single-photon exchange) is given by

$$\hat{g}_{\omega;C}(1,2) = \frac{1}{r_{12}} - \frac{\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} \exp\left(i\omega_{12} r_{12}\right) + \left(\boldsymbol{\alpha}_1 \cdot \boldsymbol{\nabla}_1\right)\left(\boldsymbol{\alpha}_2 \cdot \boldsymbol{\nabla}_2\right) \frac{\exp\left(i\omega_{12} r_{12}\right) - 1}{\omega_{12}^2 r_{12}} \tag{2.12}$$

where r_{12} is the interelectron distance. In the Feynman (Lorentz) gauge the interaction is given by

$$\hat{g}_{\omega;F}(1,2) = \frac{1 - \boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} \exp\left(i\omega_{12} r_{12}\right) \tag{2.13}$$

The exponential part $\exp\left(i\omega_{12} r_{12}\right)$ represents the retardation of the interaction. The real part of the exponential leads to shifts in energy levels, whereas the imaginary part contribute only to the level width [11] and will be omitted in this discussion (i.e. we may replace $\exp\left(i\omega_{12} r_{12}\right)$ by $\cos\left(\omega_{12} r_{12}\right)$ in the expressions above). In the expressions above we recognize $(1/r_{12})$ as the instantaneous Coulomb interaction and $\left(\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2\right)/r_{12}$ as the instantaneous magnetic interaction of the electron spins. We see that in the Lorentz gauge both the the electric and magnetic interactions are explicitly retarded, corresponding to the exchange of both transverse and longitudinal photons. In the Coulomb gauge only the instantaneous magnetic interaction is explicitly retarded, which corresponds to quantization of only the transverse part of the electromagnetic field [131, 132]. This can be understood from the discussion of Coulomb gauge above.

The momentum transfer ω of the exchanged photon refers to one-electron states such that if we consider the transition probability

$$\left|\left\langle C(1)D(2)\left|\hat{g}_\omega(1,2)\right|A(1)B(2)\right\rangle\right|^2 \tag{2.14}$$

between two-electron states $|A(1)B(2)\rangle$ and $|C(1)D(2)\rangle$, the ω is defined as

$$\omega = \omega_{AC} = \frac{|\varepsilon_A - \varepsilon_C|}{c} = \omega_{BD} = \frac{|\varepsilon_B - \varepsilon_D|}{c} \tag{2.15}$$

where the ε's are one-electron energies [127]. Consequently, the evaluation of Eq.(2.14) requires the definition of such one-electron states. This reflects the independent-particle approach inherent in QED. The quantization of the Dirac field effectively defines electrons and positrons [133, 134, 135], since the creation and annihilation operators refer to a complete set of solutions of the Dirac equation. The "free picture" employs the complete set of solutions of the free particle Dirac equation, whereas the "bound state interaction picture" or "Furry picture" [136] use the complete set of solutions of the Dirac equation in the presence of some external field. The complete set of bound state electrons and positrons are related to the set of free electrons and positrons by a unitary transformation which "dresses up" the particles.

We immediately realize that the evalution of matrix elements in configuration space involving the ω-dependent electron-electron interaction becomes exceedingly difficult. For single-configuration calculations one-electron energies may be approximated by Koopmans' theorem, whereas no clear-cut definition of one-electron energies exists in a multiconfigurational approach [127]. We will therefore have to resort to approximations which eliminates the ω-dependence. It turns out that the consequences of this approximation are not all that severe, since corrections due to the ω-dependence are found to contribute only to order $Z^5\alpha^4$ [137, 138]. In the limit $\omega \rightarrow 0$ (the low-frequency limit), the electron-electron interaction in the Coulomb gauge reduces to the Coulomb-Breit interaction

$$\lim_{\omega\rightarrow 0} \hat{g}_{\omega;C} = \hat{g}^{Coulomb} + \hat{g}^{Breit} \tag{2.16}$$

consisting of the instantaneous Coulomb interaction

$$\hat{g}^{Coulomb} = \frac{1}{r_{12}} \tag{2.17}$$

and the *Breit term* [139]

$$\hat{g}^{Breit} = -\left\{\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}} + \frac{(\boldsymbol{\alpha}_1\cdot\boldsymbol{\nabla}_1)(\boldsymbol{\alpha}_2\cdot\boldsymbol{\nabla}_2)\,r_{12}}{2}\right\} \tag{2.18}$$

where ∇_1 and ∇_2 act *only* on r_{12} and not on the wave function. The Breit term is usually expressed as

$$\hat{g}^{Breit} = -\left\{\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{2r_{12}} + \frac{(\boldsymbol{\alpha}_1\cdot\mathbf{r}_{12})(\boldsymbol{\alpha}_2\cdot\mathbf{r}_{12})}{2r_{12}^3}\right\} \tag{2.19}$$

Correspondingly, the electron-electron interaction in the Feynman gauge reduce to the Coulomb-Gaunt interaction

$$\lim_{\omega\rightarrow 0} \hat{g}_{\omega;F} = \hat{g}^{Coulomb} + \hat{g}^{Gaunt} \tag{2.20}$$

where the *Gaunt-term* [140]

$$\hat{g}^{Gaunt} = -\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}} \tag{2.21}$$

represents the direct magnetic interaction of electron spins. Note that the Breit term may be written in terms of the Gaunt term and a gauge-dependent term

$$\hat{g}^{Breit} = \hat{g}^{Gaunt} + \hat{g}^{gauge}; \qquad \hat{g}^{gauge} = \frac{(\boldsymbol{\alpha}_1\cdot\boldsymbol{\nabla}_1)(\boldsymbol{\alpha}_2\cdot\boldsymbol{\nabla}_2)\,r_{12}}{2} \tag{2.22}$$

The Coulomb-Breit interaction is correct to order $(Z\alpha)^2$. In order to achieve the same accuracy for the low-frequency limit of the electron-electron interaction in Feynman gauge, it is necessary to add a ω-dependent term

$$-\frac{1}{2}\omega^2 r_{12} \tag{2.23}$$

Due to the appearance of this term to order $(Z\alpha)^2$, it has been claimed inconsistent to use the Coulomb-Gaunt interaction in relativistic calculations [141, 130]. However, the choice of what electron-electron potential to use should be based not only on order analysis, but also on what physical effects are included and, in the final instance, on the computational demands.

Let us therefore look at the physical content of the ω-independent two-electron operators. We can do so by performing a reduction of the two-electron operators to a two-component form correct to order $(Z\alpha)^2$, analogous to the derivation of the Pauli-operator Eq.(1.163) in section 1.8. The Coulomb-Breit interaction has been reduced to two-component form by Chraplyvy et al. [142, 143, 144] by a Foldy-Wouthuysen transformation, and the result is discussed in Moss [40]. The result may be combined with the Pauli-operator to form the so-called *Breit-Pauli Hamiltonian* [145]. I have redone this derivation for the Gaunt and gauge terms separately in order to see what terms in the Breit-Pauli Hamiltonian are derived from the Gaunt term. Details about the derivation are provided in Appendix B. Here we just state results. The instantaneous Coulomb interaction reduce to

$$\begin{array}{lll} \hat{g}^{Coulomb}: & \frac{1}{r_{12}} & (a) \\ & -\dfrac{1}{4m^2c^2r_{12}^3}\left[\boldsymbol{\sigma}_1\cdot(\mathbf{r}_{12}\times\mathbf{p}_1)-\boldsymbol{\sigma}_2\cdot(\mathbf{r}_{12}\times\mathbf{p}_2)\right] & (b) \\ & -\dfrac{1}{m^2c^2}\pi\delta(\mathbf{r}_{12}) & (c) \end{array} \tag{2.24}$$

We expect the terms generated by the instantaneous Coulomb interaction to be analogous to the terms in the Pauli-operator Eq.(1.163) depending on the nuclear potential since we have simply replaced this potential by the corresponding potential from an electron. We do indeed find the expected analogy and can interpret the various terms as follows:

(a) Coulomb interaction

(b) Spin-own orbit interaction:
the spin-orbit interaction of a electron generated by the electric field of another electron.

(c) Darwin-type correction to the Coulomb term:
a correction to the Coulomb interaction Eq.(2.24a) due to *Zitterbewegung*.

The Gaunt-term reduce to

$$\begin{aligned}
\hat{g}^{Gaunt}: \quad & - \frac{1}{m^2c^2 r_{12}} (\mathbf{p}_1 \cdot \mathbf{p}_2) & (a) \\
& + \frac{1}{2m^2c^2 r_{12}^3} [\boldsymbol{\sigma}_1 \cdot (\mathbf{r}_{12} \times \mathbf{p}_2) - \boldsymbol{\sigma}_2 \cdot (\mathbf{r}_{12} \times \mathbf{p}_1)] & (b) \\
& + \frac{1}{4m^2c^2} \left[r_{12}^{-3} (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) - 3 r_{12}^{-5} (\boldsymbol{\sigma}_1 \cdot \mathbf{r}_{12}) (\boldsymbol{\sigma}_2 \cdot \mathbf{r}_{12}) \right] & (c) \\
& - \frac{2\pi}{3m^2c^2} (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) \, \delta(\mathbf{r}_{12}) & (d) \\
& - \frac{1}{2m^2c^2 r_{12}^3} (\mathbf{r}_{12} \cdot \boldsymbol{\nabla}_{12}) & (e) \\
& - \frac{1}{m^2c^2} \pi \delta(\mathbf{r}_{12}) & (f)
\end{aligned} \tag{2.25}$$

and the gauge-term contributes

$$\begin{aligned}
\hat{g}^{gauge}: \quad & - \frac{1}{2m^2c^2} (\mathbf{p}_1 \cdot \boldsymbol{\nabla}_1) (\mathbf{p}_2 \cdot \boldsymbol{\nabla}_2) \, r_{12} & (a) \\
& + \frac{1}{2m^2c^2 r_{12}^3} (\mathbf{r}_{12} \cdot \boldsymbol{\nabla}_{12}) & (b) \\
& + \frac{1}{m^2c^2} \pi \delta(\mathbf{r}_{12}) & (c)
\end{aligned} \tag{2.26}$$

By adding contributions from the two terms we obtain the reduced Breit term

$$\begin{aligned}
\hat{g}^{Breit}: \quad & - \frac{1}{m^2c^2} \left[r_{12}^{-1} (\mathbf{p}_1 \cdot \mathbf{p}_2) + \frac{1}{2} (\mathbf{p}_1 \cdot \boldsymbol{\nabla}_1) (\mathbf{p}_2 \cdot \boldsymbol{\nabla}_2) \, r_{12} \right] & (a) \\
& + \frac{1}{2m^2c^2 r_{12}^3} [\boldsymbol{\sigma}_1 \cdot (\mathbf{r}_{12} \times \mathbf{p}_2) - \boldsymbol{\sigma}_2 \cdot (\mathbf{r}_{12} \times \mathbf{p}_1)] & (b) \\
& + \frac{1}{4m^2c^2} \left[r_{12}^{-3} (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) - 3 r_{12}^{-5} (\boldsymbol{\sigma}_1 \cdot \mathbf{r}_{12}) (\boldsymbol{\sigma}_2 \cdot \mathbf{r}_{12}) \right] & (d) \\
& - \frac{2\pi}{3m^2c^2} (\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) \, \delta(\mathbf{r}_{12}) & (e)
\end{aligned} \tag{2.27}$$

where the individual terms are interpreted as folllows

(a) Orbit-orbit interaction:

This term corresponds exactly to the Breit term Eq.(2.18) if we make the substitution

$$\boldsymbol{\alpha} \rightarrow \frac{\mathbf{p}}{mc} \tag{2.28}$$

This corresponds to the transition from the relativistic velocity operator $c\boldsymbol{\alpha}$ Eq.(1.171) to the corresponding non-relativistic operator $(\mathbf{p}/m)$. The orbit-orbit term represents the classical retardation correction to the Coulomb interaction in the Coulomb gauge. It is readily rearranged to its more conventional form [146]

$$= -\frac{1}{m^2c^2}\left\{\frac{\mathbf{p}_1\cdot\mathbf{p}_2}{2r_{12}} + \frac{(\mathbf{p}_1\cdot\mathbf{r}_{12})(\mathbf{p}_2\cdot\mathbf{r}_{12})}{2r_{12}^3}\right\} \tag{2.29}$$

The first term, which is derived from the Gaunt term, may be interpreted as a current-current interaction. The second term stems from $\hat{g}^{gauge}$ and is consequently gauge-dependent.

(b) Spin-other orbit interaction:
the interaction of the spin of an electron with the magnetic field generated by the current of another electron. There is no analogous term arising from the molecular field, since teh nuclei are assumed to be a stationary sources.

(c) Spin-spin interaction:
direct dipolar interaction between two spins.

(d) Fermi-contact interaction:
it can be regarded as a finite-size correction to the spin-spin interaction

In conclusion, we see that with the exception of the gauge-dependent term Eq.(2.26a), which contributes to the orbit-orbit interaction, the whole physical content of the Breit term is provided by the Gaunt term as well. The Gaunt term in addition gives rise to two terms. The first Eq.(2.25f) corresponds to a Darwin-type correction of the Coulomb interaction, whereas the other has a less straightforward physical interpretation. The two extra terms are cancelled by corresponding terms from $\hat{g}^{gauge}$. It should be noted that none of the extra terms are spin-dependent, which means that the Coulomb-Gaunt interaction will give the total spin-orbit interaction correct to order $(Z\alpha)^2$. In general, the Gaunt term accounts for about 90% of the Breit term in atomic calculations and shifts the total energy upwards, whereas the gauge term lowers the total energy [147]. Multiconfigurational calculations on helium-like ions [127] has shown that at around $Z = 50$ the correlation energy from the Gaunt term becomes larger than the correlation from the Coulomb term. However, the Gaunt term appears to be of importance only near nuclei [126], which makes the Gaunt interaction mainly localized to atoms, so that it is not expected to significantly change molecular properties such as bond lengths. Our Dirac-Fock calculations on PtH (paper IV) and previous calculations on hydrides of group IVA by Visser et al.[148] show that the Gaunt term has a negligible effect on molecular bond lengths. Visser et al. [148] found that the Gaunt term favors bond expansion, yet in PbH_4 the expansion is only 0.17 pm.

From the discussion above, we can conclude as follows: The Coulomb interaction, which may be regarded as the zeroth order term of an expansion in $(Z\alpha)^2$ of the full electron-electron interaction, appears to be sufficient for the calculation of molecular potential surfaces. Corrections of order $(Z\alpha)^2$ to the Coulomb-interaction are largely atomic in nature and need only be considered when properties dependent on the electronic density in the nuclear region are studied. The Gaunt term alone ensures correct contributions from electronic potentials to the total spin-orbit interaction to order $(Z\alpha)^2$ and may therefore be included in accurate calculations of spin-orbit splittings. The Gaunt term is straightforwardly implemented in a finite basis approximation, since it reduces to ordinary Coulomb repulsion integrals in a scalar basis. The gauge-dependent term, on the other hand, leads to more complicated two-electron integrals.

2.2 Relativistic molecular calculations

> *There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact.*
>
> M.TWAIN *(1874) [149]*
>
> *Theoreticians get such amazing returns from such a small investment of fact combined with a few hours on a Cray Y/MP.*
>
> I.M.KLOTZ *(1990)*

2.2.1 The Dirac-Coulomb Hamiltonian

In this section we arrive at the main goal of our study, namely relativistic molecular calculations. Ideally we would like a Lorentz invariant description of molecular systems, but from the discussion in the previous sections, it is clear that such a description can only be approximate. Let us briefly review the approximations involved.

First, the Born-Oppenheimer approximation is inherently incompatible with the theory of special relativity since it singles out a preferred reference frame, namely the frame in which nuclei can be treated as stationary sources of external fields. Relativistic corrections to the nuclear motion are, however, expected to be small [150]. The advantage of the Born-Oppenheimer approximation is that it reduces the complexity of the molecular description and allows us to focus our attention on the electronic degrees of freedom. The eigenvalues of the electronic Hamiltonian defined by the Born-Oppenheimer approximation are assumed to vary smoothly as a function of nuclear coordinates. This leads to the concept of molecular potential energy surfaces.

Second, we have neglected all hyperfine effects, that is we have neglected nuclear spins. As discussed in section 1.7, this leads to a time reversal symmetric Hamiltonian. The hyperfine effects also include effects of a possible nuclear electric quadrupole moment. We can model this, for example by inclusion of Gaussian d functions in a finite nucleus model.

Third, the description of even a single electron in the molecular field leads to a many-body problem due to the possible creation of virtual electron-positron pairs. The proper treatment of the problem can only be obtained within the framework of QED, which allows the number of particles in the system to vary. We avoid working with the full mathematical machinery of QED by invoking the no-pair approximation, that is we neglect all pair creations. This means that we stay within the framework of Dirac's hole theory with the Dirac sea of negative energy electrons at all times completely filled. It corresponds to working with classical fields and implies neglect of QED effects, such as self energy and vacuum polarization, which represent the interaction of the electron with the zero-point fluctuations of the quantized electromagnetic and Dirac fields, respectively [87].

Fourth, the electron-electron interaction can be handled correctly only to order $(Z\alpha)^2$

if we want to avoid ω-dependent terms. Note that this approximation is closely tied to the no-pair approximation. If we allow virtual pair creation, we must take into account electron-positron interaction in terms of virtual photon exchange, and this leads to non-negligible ω-terms of order 2mc. The level of accuracy obtained with the truncated electron-electron interaction is, however, expected to be sufficient for all relevant chemical properties. In fact, for the determination of molecular potential energy surfaces, experience so far [148, 31] indicate that the Coulomb interaction alone appears to provide an adequate description.

We therefore choose the *Dirac-Coulomb Hamiltonian* as the starting point for relativistic molecular calculations, but may consider inclusion of the Gaunt-term for the study of properties that depend on the electron density near the nuclei, or when we want very accurate spin-orbit splittings. The Dirac-Coulomb Hamiltonian for a molecular system of n electrons in the field of N nuclei has the form

$$\hat{H}_{DC} = \sum_{i=1}^{n} \hat{h}_{D;V}(i) + \sum_{i<j}^{n} \hat{g}^{Coulomb}(i,j) + \hat{V}_{N-N} \tag{2.30}$$

The first term is a sum over one-electron Dirac operators in the molecular field (see section 1.7)

$$\hat{h}_{D;V} = \beta' mc^2 + c\,(\boldsymbol{\alpha}\cdot\mathbf{p}) + \hat{V}_{e-N} \tag{2.31}$$

The second term describes the electron-electron interaction in terms of the instantaneous Coulomb interaction. Even though this operator has the same operator form as the electron-electron interaction in non-relativistic theory, its physical content is different since it includes spin-own orbit interactions and Darwin-type corrections to the Coulomb interaction, as seen in section 2.1. The last term is the Coulomb interaction of nuclei.

$$\hat{V}_{N-N} = \sum_{I<J}^{N} \frac{Z_I Z_J}{r_{IJ}} \tag{2.32}$$

The Dirac-Coulomb Hamiltonian is an intuitive extension of the non-relativistic electronic Hamiltonian , but the validity of $\hat{H}_{DC}$ and the resulting Dirac-Coulomb equation

$$\hat{H}_{DC}\Psi = E\Psi \tag{2.33}$$

has been surrounded by considerable controversy. We shall return to this in section 2.2.6. Let us first note, however, that the inclusion of the electron-electron interaction enforces a fifth approximation in our description of molecular systems, since the Dirac-Coulomb equation represents a many-body problem and has no analytical solutions. From a mathematician's point of view the Dirac-Coulomb equation is nightmarish, since it constitutes a set of multivariable coupled first order differential equations with one singularity for every pair of particles in the system. Yet we can provide approximate solutions to it of high accuracy. The key to success lies in the use of physical-chemical understanding in the construction of approximate solutions.

2.2.2 The variational method

Approximate solutions to the Dirac-Coulomb equation can be obtained by perturbation theory or by the variational method. We shall consider the latter method. The basic idea of the variational method is to introduce a trial function furnished with parameters that can be varied so as to obtain the best possible approximate solution within the parameter space. The parametrization of the trial function leads to a parametrization of its energy, defined as the expectation value of the Dirac-Coulomb Hamiltonian. Approximations to the exact eigenfunctions of the Hamiltonian are found as stationary values of the energy in the parameter space [151]. Note that if the variational parameters are introduced in a non-linear manner, the reciprocal relation does not hold true, so that a stationary value of the energy may correspond to a physically unacceptable solution of the Dirac-Coulomb equation [152]. Let us consider the general form of the trial function in molecular electronic structure theory.

The basic building blocks for approximative wave functions are molecular orbitals (MOs) and electronic configurations. They can be introduced as follows: If we turn off the electron-electron interaction, the electronic Hamiltonian reduce to a sum of one-electron Dirac operators, and the wave function may be written as a Hartree product of one-electron molecular 4-spinors

$$\Phi = \prod_{i=1}^{n} \psi_i(\mathbf{r}_i) \tag{2.34}$$

The spinors are chosen from the complete set $\{\psi_i\}$ of orthonormal solutions to the corresponding Dirac equation in the molecular field. We shall refer to any set of one-electron functions as our 1-particle basis and denote the individual one-electron functions molecular orbitals (MO). Physically Eq.(2.34) is, however, not an acceptable many-electron wave function, since it does not obey the Pauli-principle, which states that the many-electron wave function should change sign under the permutation of any pair of electrons (fermions). We can remedy the situation by antisymmetrizing the wave function, for example by writing it as a Slater-determinant

$$\Phi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \dots & \psi_n(1) \\ \psi_1(2) & \psi_2(2) & \dots & \psi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(n) & \psi_2(n) & \dots & \psi_n(n) \end{vmatrix} \tag{2.35}$$

of orthonormal molecular orbitals

$$\langle \psi_i \mid \psi_j \rangle = \delta_{ij} \tag{2.36}$$

The Slater-determinant represents a particular electronic configuration, namely the set of molecular orbitals appearing in the determinant. When the electron-electron interaction is turned on, it is no longer possible to separate the electronic degrees of freedom, and the exact electronic wave function can no longer be expressed in terms of a single Slater-determinant. It can, however, be expanded in all possible Slater-determinants that can be generated from the complete 1-particle basis:

$$\Psi = \sum_i c_i \Phi_i \tag{2.37}$$

The exact solution is thereby defined in terms of complete 1- and n-particle bases, where the set $\{\Phi_i\}$ of Slater-determinants in Eq.(2.37) constitute the n-particle basis. We may therefore seek approximate solutions of the Dirac-Coulomb equation in the space of truncated 1 and n-particle bases. The search for stationary energies then corresponds to separate rotations within the 1-particle basis and within the n-particle basis.

The simplest variational approach is to (in the closed-shell case) choose a single Slater determinant Eq.(2.35) as the trial function. This forms the basis for the Hartree-Fock (HF) method in non-relativistic theory and the Dirac-Fock (DF) method in relativistic theory. These are independent particle models that view the electron as moving independently in the field of the nuclei and the average field of the other electrons. The independent particle model usually provides an adequate description of molecular structure at the equilibrium geometry, but fails in situations where degeneracies or near-degeneracies of configurations occur. Near-degeneracies typically arise in bond breaking and bond formation, in open-shell and excited states. In relativistic molecules, additional near-degeneracies may be introduced by the fine structure of the spin-orbit splitting. Such systems require a multiconfigurational approach. The most compact and flexible description is provided by the Multi-Configurational Self-Consistent Field (MCSCF) method, which allows both configurational and orbital parameters to vary.

At this point, it will be convenient to summarize the papers included in the thesis in order to avoid overlap in the presentation.

2.2.3 Summary of papers

- PAPER I:

T. Saue, T.Helgaker, K. Fægri and O.Gropen:
Principles of direct 4-component relativistic SCF:
Application to dihydrides of group 16
Manuscript.

In this paper we present the theory of direct 4-component SCF calculations based on a quaternion formulation of the Dirac-Fock equations. In the closed-shell Dirac-Fock model, bound electronic states of the Dirac-Coulomb Hamiltonian Eq.(2.30) are approximated by a trial function in the form of a single Slater-determinant Eq.(2.35). In the finite basis approximation we expand each component of the molecular 4-spinors in a real (AO) basis We seek stationary values of the total energy under the constraint of orthonormal molecular orbitals. As in the non-relativistic theory, this leads to a pseudoeigenvalue equation in terms of the expansion coefficients $\mathbf{c}$. We proceed by showing that due to time reversal symmetry we may reduce the Dirac-Fock equations to a quaternion form Eq.(1.101)

$$ {}^{Q}\mathbf{F}\,{}^{Q}\mathbf{c} = \left[\mathbf{F}^{\alpha\alpha} + \mathbf{F}^{\alpha\beta}\check{j}\right]\left[\mathbf{c}^{\alpha} - \mathbf{c}^{\beta *}\check{j}\right] = \varepsilon\,{}^{Q}\mathbf{S}\left[\mathbf{c}^{\alpha} - \mathbf{c}^{\beta *}\check{j}\right] = \varepsilon\,{}^{Q}\mathbf{S}\,{}^{Q}\mathbf{c} \tag{2.38} $$

This reduces the operation count and memory requirement for the construction of the Fock matrix by a factor two. The quaternion formalism furthermore brings the two-electron Fock matrix onto a form that is readily incorporated into existing software for non-relativistic calculations. By a quaternion diagonalization of the Fock matrix we obtain well-defined Kramers partners, which allows efficient use of time reversal symmetry in post-DF applications.

The quaternion Dirac-Fock equations are solved iteratively by the direct SCF method, in which two-electron integrals are regenerated in each SCF iteration. This eliminates the frequent problems with disk storage and I/O load in the standard Dirac-Fock approach and makes DF-calculations on workstations feasible. Integral batches are prescreened based on the differential density matrix approach. The integral presecreening is supplemented by separate screening of Coulomb and exchange contributions to the Fock matrix. The SCF convergence is accelerated by the implementation of the DIIS method.

We have applied the 4-component direct SCF method to the dihydrides of tellurium, polonium and eka-polonium (element 116). We find the expected bond shortening due to relativity in H_2Te and H_2Po. In the dihydride of element 116 we observe, however, a dramatic bond expansion due to the extreme spin-orbit splitting of the 7p orbital in the eka-polonium atom. This spin-orbit effect is further analyzed in paper V.

- PAPER II:

T. Saue and H.J.Aa. Jensen:

Quaternion symmetry in relativistic molecular calculations:
I. The Dirac-Fock method

Submitted to J. Chem. Phys.

In this paper we consider the use of the full symmetry group of the Dirac-Coulomb Hamiltonian to reduce the computational effort in the Dirac-Fock method. The full symmetry group consists of all operators, unitary or antiunitary, that commute with the Dirac-Coulomb Hamiltonian. In practice, we obtain the full symmetry group as the direct product of the time reversal operator and the molecular point group. Due to the presence of antiunitary operators we can no longer form true matrix representations of the group. It is, however, still possible to establish a system of matrices, a corepresentation that can be broken down to irreducible forms. We denote the irreducible forms ircops.

We limit spatial symmetry to D_{2h} and subgroups, which we collectively denote binary groups, since they constitute the set of all single point groups with no elements of order higher than two. The binary groups can be classified as quaternion, complex and real based on the distribution of Kramers partners among the fermion irreps of the molecular double point group:

1. Quaternion groups: C_1, C_i
2. Complex groups: C_s, C_2, C_{2h}
3. Real groups: C_{2v}, D_2, D_{2h}

In the case of real (complex) groups the Fock matrix in a Kramers restricted MO-basis automatically reduce to a real (complex) matrix. It is, however, the construction of the Fock matrix in AO-basis that constitute the time-consuming step in a Dirac-Fock calculation. We show that by a simple quaternion transformation of the real basis, it is possible to obtain the same matrix reduction in the AO-basis as well. The symmetry scheme amounts to a simple scheme of phase insertion that require virtually no extra computational effort, but leads to considerable computational gains, as is demonstrated by a numerical example.

- PAPER III:

H.J.Aa.Jensen, K.G.Dyall, T.Saue and K.Fægri:
Relativistic 4-component Multi-Configurational Self-Consistent Field Theory for Molecules: Formalism
Accepted for publication in J.Chem.Phys.

In this paper we outline the most general multi-configurational approach to relativistic molecular calculations, namely the multi-configurational self-consistent field (MCSCF) method in which both configurational and orbital parameters are allowed to vary. We employ the second quantization formalism whereby the Dirac-Coulomb Hamiltonian Eq.(2.30) is expressed by

$$\hat{H}_{DC} = \sum_{pq} \left\langle p \left| \hat{h}_{D;V} \right| q \right\rangle p^{\dagger} q + \frac{1}{2} \sum_{pqrs} \left\langle pq \left| \hat{g}^{Coulomb} \right| rs \right\rangle p^{\dagger} r^{\dagger} sq \tag{2.39}$$

where $p^{\dagger}$ and p are creation and annihilation operators for 4-spinor p. Time reversal symmetry is enforced by transformation to Kramers restricted basis. This leads to the introduction of Kramers single $X^{\pm}_{pq}$ and double $x^{\pm,\pm}_{pq,rs}$ replacement operators.

The parametrized Kramers restricted MCSCF wave function has the form

$$|MC(\boldsymbol{\delta}, \boldsymbol{\kappa})\rangle = \exp\left(-\hat{\kappa}\right) |\delta\rangle \tag{2.40}$$

where $\exp\left(-\hat{\kappa}\right)$ is an exponential parametrization of orbital rotations and where the configurational variational parameters $\boldsymbol{\delta}$ describe a correction vector orthogonal to the current configuration expansion.

MCSCF methods require second order optimization methods for controllable convergence. This implies knowledge of the Hessian (second derivative) matrix (or approximations to it). Key ingredients in the formalism that make large configuration expansions feasible are:

1. The unitary parametrization $\exp\left(-\hat{\kappa}\right)$ of the orbital optimization ensures the orthonormality of molecular orbitals, so that unconstrained optimization techniques can be used.

2. The Hessian matrix times a vector is calculated directly by iterative techniques so that the individual elements of the matrix need not be known.

3. The restricted step second order optimization techniques is a robust technique for sharp and well-controllable convergence in relatively few iterations.

A major difference with the corresponding non-relativistic method is the replacement of the non-relativistic minimization principle with a minimax principle for ground state optimization.

- PAPER IV:

L.Visscher, T.Saue, W.C.Nieuwpoort, K.Fægri and O.Gropen:
The electronic structure of the PtH molecule:
Fully relativistic configuration interaction calculations
of the ground and excited states
J.Chem.Phys. **99** (1993) 6704–6715

In this paper we investigate the lower five states of the the open-shell molecule platinum hydride by 4-component configuration interaction (CI) calculations. We have used the MOLFDIR [22] package for relativistic molecular calculations. The CI program is based on the restricted active space (RAS) formalism. The desired roots of the CI Hamiltonian matrix are found by iterative techniques.

The characteristics of the five lower states are largely dominated by the spin-orbit splitting of the 5d orbital in the platinum atom, which is of the order 10 000 cm^{-1}. Thus we find three lower states arising from the splitting of a $5\mathrm{d}^4_{3/2}5\mathrm{d}^5_{5/2}\sigma^2_{1/2}$ electronic configuration in the molecular field and two upper states arising from the $5\mathrm{d}^3_{3/2}5\mathrm{d}^6_{5/2}\sigma^2_{1/2}$ configuration. The bonding is to a large extent a σ(s–s) bond, but with some contribution from the platinum 5d orbitals. The effect of the Gaunt term was investigated at the SCF level by including it perturbatively, and was found to be negligible for spectroscopic properties of PtH. The main correlation effects stem from the angular correlation of the 5d orbitals and lead to bond shortages on the order 4 pm and an increase in the dissociation energy of about 0.5 eV. The high stability of the Pt-H bond can be explained by the relativistic stabilization of the 6s orbital.

- PAPER V:

T.Saue, K.Fægri and O.Gropen:

Relativistic effects on the bonding of heavy and superheavy hydrogen halides

Submitted to Phys.Rev.Lett.

In this paper we have investigated bonding in heavy and superheavy hydrogen halides by direct 4-component SCF calculations using the DIRAC code. We find a relativistic bond contraction of 0.6 pm in hydrogen iodide. In hydrogen astatide there is, however, a slight bond expansion of 0.3 pm, and in the hydride of eka-atstatine (element 117) we find a huge bond expansion of 12.9 pm.

We have analyzed the bonding in the three molecules by projecting the molecular orbitals down onto the vectors of the halide ions. It then becomes clear that the huge bond expansion observed for the hydride of eka-astatine is due to the extreme spin-orbit splitting of the atomic 7p orbital. The bonding in the corresponding hydride is thereby dominated by the radially diffuse $7p_{3/2}$ orbital.

2.2.4 The small component density

The small components are denoted "small" because they tend towards zero for electronic solutions in the non-relativistic limit $c \rightarrow \infty$, as discussed in section 1.8. In this section I consider the smallness of the small components for the finite speed of light. I have calculated the small component density for the elements $Z = 1$ to $Z = 103$ using the numerical 4-component atomic code GRASP [117]. The results are presented in Tab.2.2.4 and graphically in Fig.2.1. It is seen that the small components contribute an electron only for the heaviest element $Z = 103$ in the series. Also, we note from Fig.2.1 that there is no sign of periodic structure in the small component density as a function of atomic number. Rather we see a smooth curve that is readily fitted by a low-order polynomial of the atomic number. This indicates that the small component density is largely located to the nuclear region, where it experiences the almost unscreened nuclear charge. I have investigated this further by plotting the small component density for the radon atom $Z = 86$ in Fig.2.2. It shows that the small component density is approximately limited to a region within 0.2 bohrs from the nucleus. For comparison, the radial expectation values of the atomic spinors with $n = 2$ are found in the region 0.1 – 0.2 bohrs. In Tab.2.2.4 the accumulation of the small component density is listed for the radon atom. It shows that the outer shell ($n = 6$) of the radon atom accounts for only 0.13 % of the total small component density of 0.62805 electrons.

The small component density is seen to be highly localized and thereby atomic in nature, so that we do not expect the appreciable changes in the small component density when the atom enters a molecule. This suggests that the interaction of small component densities at different atomic centers may be modelled by Coulombic repulsion. Visscher [153] found that the neglect of SS integrals in a relativistic calculation on At_2 using the 4-component version of coupled-cluster singles and doubles with perturbative treatment of triples CCSD(T) [32] led to an error in the bond length of 10.8 pm compared to a calculation with all integrals included. He was, however, able to correct this error completely by representing the contribution from the SS integrals by a Coulombic interaction of point charges Eq.2.32 using the small component density of the astatine atom. For hydrides of heavy atoms, the neglect of SS integrals appears to have negligible influence on bond lengths [154, 155], but this is due to the fact that the small component density for the hydrogen atom is for all purposes equal to zero, so that there is no Coulombic repulsion from the SS integrals.

Table 2.1: Total small component density for all elements from Z = 1 to Z = 103. The numbers marked with an asterisk have been obtained by interpolation.

1	H	0.0000	22	Ti	0.0227	43	Tc	0.1141	64	Gd	0.3003∗	85	At	0.6097
2	He	0.0001	23	V	0.0252	44	Ru	0.1206	65	Tb	0.3118	86	Rn	0.6281
3	Li	0.0002	24	Cr	0.0279	45	Pd	0.1273∗	66	Dy	0.3238	87	Fr	0.6468
4	Be	0.0004	25	Mn	0.0308	46	Pd	0.1343	67	Ho	0.3360	88	Ra	0.6659
5	B	0.0007	26	Fe	0.0339	47	Ag	0.1415	68	Er	0.3485	89	Ac	0.6854
6	C	0.0010	27	Co	0.0371	48	Cd	0.1489	69	Tm	0.3614	90	Th	0.7053
7	N	0.0014	28	Ni	0.0405	49	In	0.1566	70	Yb	0.3746∗	91	Pa	0.7256
8	O	0.0020	29	Cu	0.0440	50	Sn	0.1644	71	Lu	0.3879	92	U	0.7464∗
9	F	0.0026	30	Zn	0.0478	51	Sb	0.1725	72	Hf	0.4017	93	Np	0.7675∗
10	Ne	0.0034	31	Ga	0.0517	52	Te	0.1809	73	Ta	0.4157	94	Pu	0.7888
11	Na	0.0043	32	Ge	0.0558	53	I	0.1895	74	W	0.4301	95	Am	0.8107
12	Mg	0.0053	33	As	0.0602	54	Xe	0.1983	75	Re	0.4447	96	Cm	0.8332∗
13	Al	0.0065	34	Se	0.0647	55	Cs	0.2073	76	Os	0.4597	97	Bk	0.8560∗
14	Si	0.0077	35	Br	0.0694	56	Ba	0.2166	77	Ir	0.4750	98	Cf	0.8790
15	P	0.0091	36	Kr	0.0743	57	La	0.2261	78	Pt	0.4907	99	Es	0.9027
16	S	0.0106	37	Rb	0.0793	58	Ce	0.2359∗	79	Au	0.5067	100	Fm	0.9268
17	Cl	0.0123	38	Sr	0.0846	59	Pr	0.2459	80	Hg	0.5230	101	Md	0.9514
18	Ar	0.0141	39	Y	0.0901	60	Nd	0.2563	81	Tl	0.5396	102	No	0.9765
19	K	0.0160	40	Zr	0.0958	61	Pm	0.2668	82	Pb	0.5566	103	Lr	1.0020
20	Ca	0.0181	41	Nb	0.1017	62	Sm	0.2777∗	83	Bi	0.5739			
21	Sc	0.0203	42	Mo	0.1078	63	Eu	0.2888	84	Po	0.5916			

Table 2.2: Accumulation of small component density in the radon atom.

	Individual contributions	Accumulated density	
$1s_{1/2}$	.1089	.21789	(34.69%)
$2s_{1/2}$	$.2537 \cdot 10^{-1}$	.26864	(42.77%)
$2p_{1/2}$	$.2544 \cdot 10^{-1}$	.31952	(50.87%)
$2p_{3/2}$	$.2206 \cdot 10^{-1}$	.40777	(64.93%)
$3s_{1/2}$	$.8629 \cdot 10^{-2}$	.42503	(67.67%)
$3p_{1/2}$	$.8564 \cdot 10^{-2}$	.44216	(70.40%)
$3p_{3/2}$	$.7592 \cdot 10^{-2}$	.47253	(75.24%)
$3d_{3/2}$	$.7449 \cdot 10^{-2}$	.50231	(79.98%)
$3d_{5/2}$	$.7191 \cdot 10^{-2}$	.54546	(86.85%)
$4s_{1/2}$	$.3011 \cdot 10^{-2}$	.55148	(87.81%)
$4p_{1/2}$	$.2913 \cdot 10^{-2}$	.55731	(88.74%)
$4p_{3/2}$	$.2568 \cdot 10^{-2}$	.56758	(90.37%)
$4d_{3/2}$	$.2355 \cdot 10^{-2}$	.57700	(91.87%)
$4d_{5/2}$	$.2261 \cdot 10^{-2}$	.59057	(94.03%)
$4f_{5/2}$	$.1897 \cdot 10^{-2}$	.60195	(95.84%)
$4f_{7/2}$	$.1858 \cdot 10^{-2}$	.61682	(98.21%)
$5s_{1/2}$	$.8433 \cdot 10^{-3}$	.61850	(98.48%)
$5p_{1/2}$	$.7607 \cdot 10^{-3}$	.62002	(98.72%)
$5p_{3/2}$	$.6503 \cdot 10^{-3}$	.62262	(99.14%)
$5d_{3/2}$	$.4733 \cdot 10^{-3}$	.62452	(99.44%)
$5d_{5/2}$	$.4456 \cdot 10^{-3}$	.62719	(99.86%)
$6s_{1/2}$	$.1531 \cdot 10^{-3}$	.62750	(99.91%)
$6p_{1/2}$	$.1118 \cdot 10^{-3}$	.62772	(99.95%)
$6p_{3/2}$	$.8238 \cdot 10^{-4}$	.62805	(100.00%)

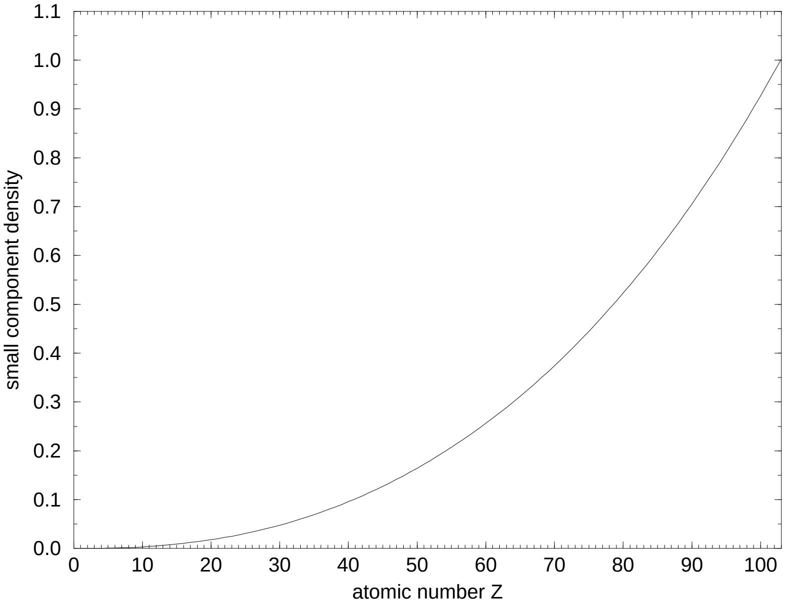


Figure 2.1: The total small component density as a function of atomic number.

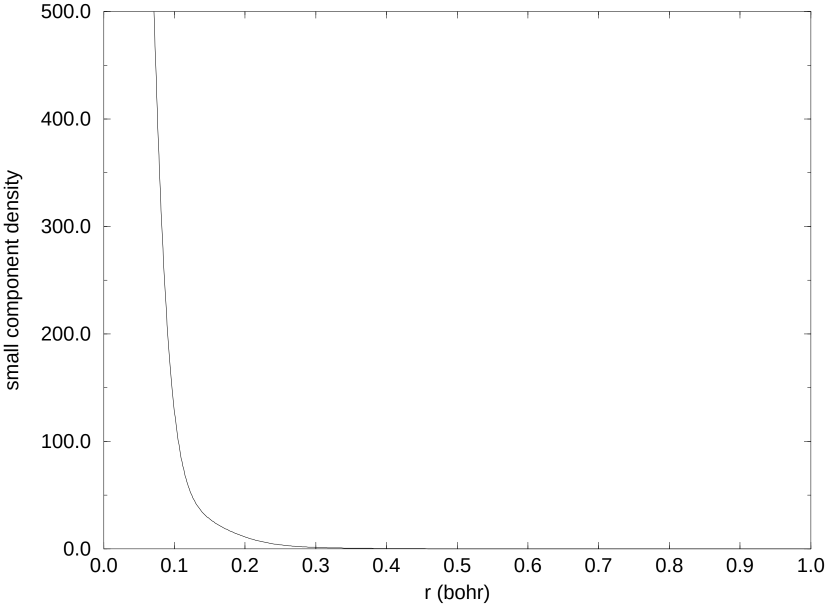


Figure 2.2: Small component density of the radon atom (Z = 86).

2.2.5 "Unphysical" solutions

In paper I we have discussed basis sets in relativistic molecular calculations. Due to their mutual coupling Eq.1.130, the large and small components have to be expanded in separate basis sets. It is customary to relate the large $\left\{\chi^L\right\}$ and small $\left\{\chi^S\right\}$ basis sets by the kinetic balance condition [26, 27]

$$\left\{\chi^S\right\} = \left\{(\boldsymbol{\sigma}\cdot\mathbf{p})\,\chi^L\right\} \tag{2.41}$$

which represents the non-relativistic limit of the exact coupling. In scalar basis sets, the kinetic balance prescription in general leads to a small component basis set that is about twice the size of the corresponding large component basis set. Charge conjugation symmetry (section 1.6) leads to a one-to-one matching of electronic and positronic solutions of the free particle Dirac equation. Obviously such a matching is not possible in a finite basis approximation to the free particle equation if the large and small component bases are of different sizes. Accordingly, the finite basis solutions to the free particle Dirac equation will contain positronic solutions with no electronic counterpart. From charge conjugation symmetry it follows that they must have eigenvalues of exactly $-2mc^2$ ($= -37557.75$ Hartrees) in the non-relativistic energy scale. They have no contribution from the large components and have therefore been characterized as unphysical. In this section we show that the "unphysical" solutions have a very physical behaviour which sheds light on the physics of positrons and on the Dirac-Fock method itself. We shall furthermore show how they can be deleted from the molecular orbital (MO) space by a simple method to obtain a one-to-one matching of basis sets.

The "unphysical" solutions have only rest mass and no kinetic energy. From the uncertainity principle they must therefore be delocalized over all space, to the extent that this is possible in the finite basis approximation. Consider now the behaviour of the "unphysical" solutions in the presence of an atom. Due to their extreme delocalization, the "unphysical" solutions see no atomic structure, only a point in space with charge equal to the total charge of the atom. A positive charge is repulsive and will therefore only introduce a downward shift in the energy of the positron. A negative charge will be attractive and introduce bound positronic states.

I have explored this picture by a series of finite basis calculations. First I solved the Dirac equation for hydrogenlike atoms of variable nuclear charges in a large uncontracted Cartesian Gaussian basis. The large component basis (22s17p14d6f) consisted of 217 functions, and the small component basis (17s36p23d14f6g), generated by unrestricted kinetic balance (see paper I), consisted of 493 functions. In Fig.2.3 I have plotted the eigenvalue of the upper positronic solution as a function of nuclear charge. A perfect linear fit is observed, with an intercept at Z = 0 at approximately $-2mc^2$. The result is therefore in complete agreement with the picture outlined above.

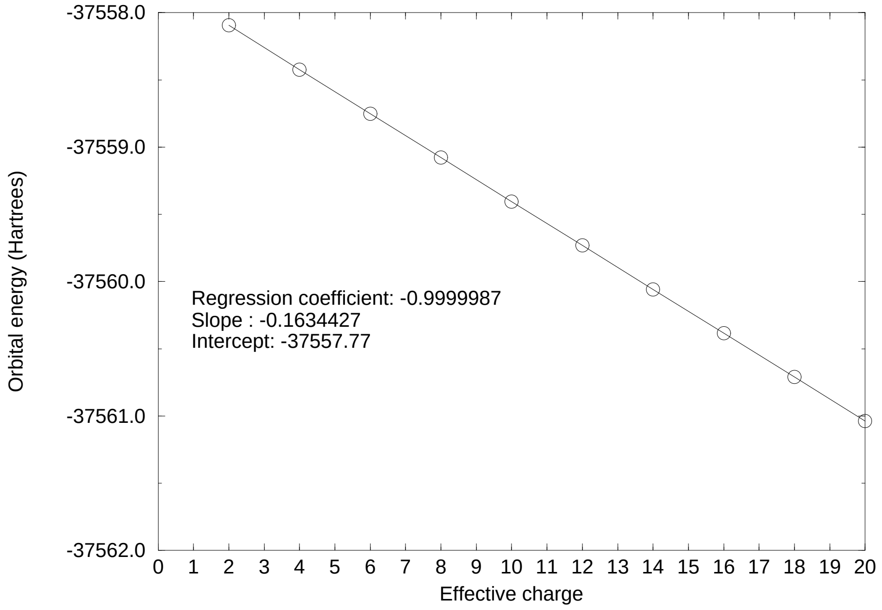


Figure 2.3: The orbital energy of the upper positronic finite basis solution to the Dirac equation for hydrogen-like atoms plotted as a function of nuclear charge Z.

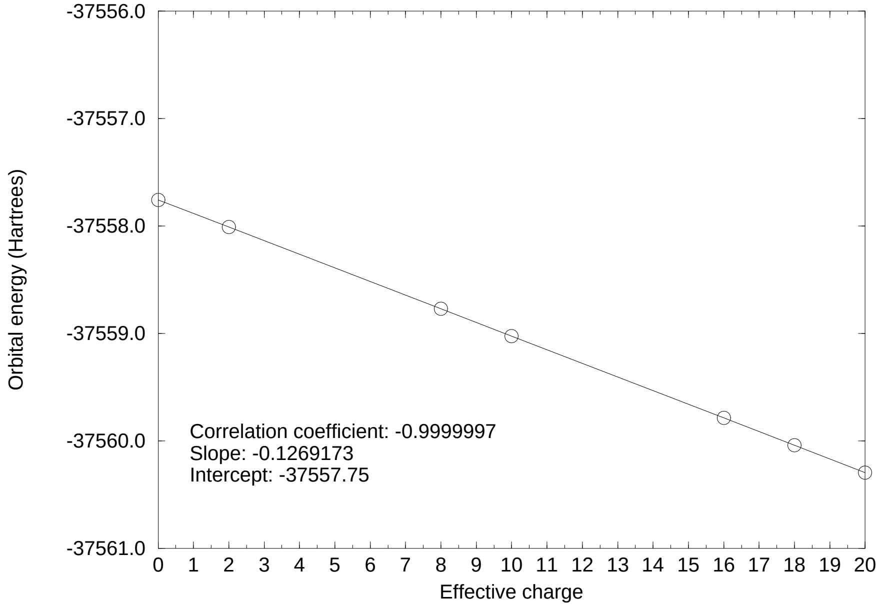


Figure 2.4: The orbital energy of the upper positronic finite basis Dirac-Fock solutions for the calsium atom where the total charge of the system is varied by changing the number of the electrons.

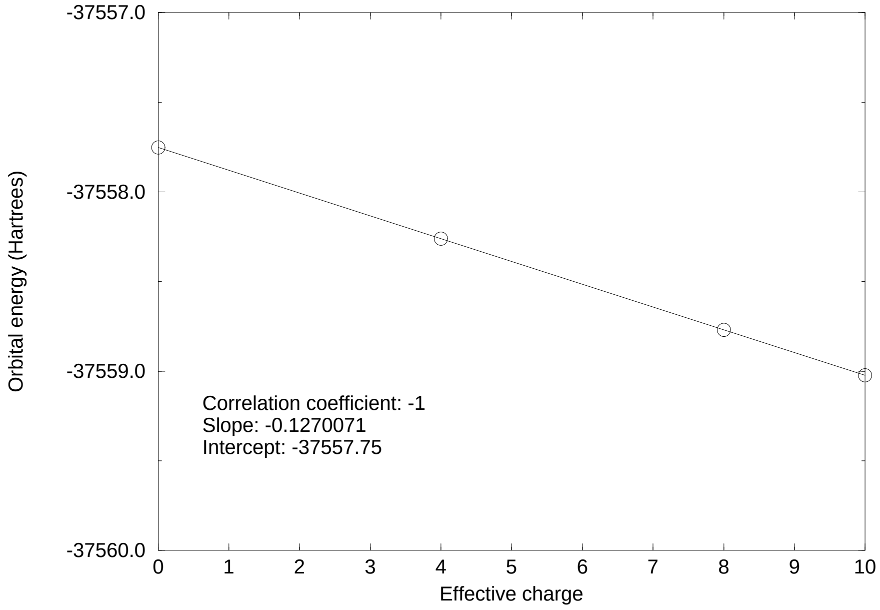


Figure 2.5: The orbital energy of the upper positronic finite basis Dirac-Fock solutions for 10-electron systems where the total charge of the system is varied by varying the nuclear charge Z.

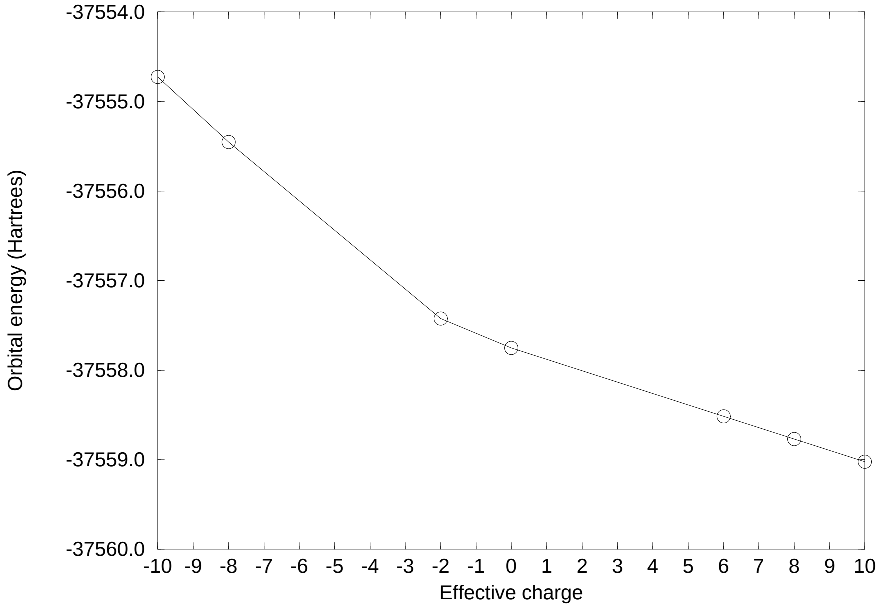


Figure 2.6: The orbital energy of the upper positronic finite basis Dirac-Fock solutions for the neon atom where the total charge of the system is varied by changing the number of electrons.

In order to test the sensitivity of the "unphysical" positronic eigenvalues to atomic structure, I performed three series of Dirac-Fock calculations in an uncontracted Cartesian Gaussian basis. The large component basis (16s11p) consisted of 49 functions and the small component basis (11s16p11d), generated by unrestricted kinetic balance, consisted of 114 functions. The first series of calculations were performed on the calsium-atom ($Z = 20$) with variable electron occupation so as to vary the total charge of the system. The upper positronic eigenvalue is plotted as a function of total charge in Fig.2.4. Again we observe a linear fit, which clearly demonstrates that the delocalized positronic solution effectively sees the calcium atom as a single point charge.

An important point regarding the Dirac-Fock (and Hartree-Fock) method is seen from Fig.2.4. The total charge of the system is the sum of the nuclear charge and the charges of the electrons in the system. A one-electron calcium atom accordingly has a total charge of +19. We solve this system by diagonalising the matrix representation of the one-electron Dirac Hamiltonian in the presence of the external field represented by a nucleus of $Z = 20$. It turns out, however, that the upper positronic solution has an eigenvalue that from the linear fit in Fig.2.4 correspond to an effective charge of +20 and not +19. We can understand how this comes about, since the one-electron Dirac Hamiltonian does not contain any potential term stemming from a single electron. From the point of view of the Dirac-Fock method, it represents a zero-electron system. In the Dirac-Fock equations for an n-electron system, the occupied Dirac-Fock orbitals describe electrons moving in the average field of the $n-1$ other electrons. On the other hand, the virtual orbitals, electronic or positronic, experience the average potential of all n electrons [3]. The question of how to obtain the positronic solutions to the system corresponding to total charge +19 is food for thoughts.

In order to further test the insensitivity of the "unphysical" solutions to the atomic structure, I fixed the number of electrons in the system to ten and then varied the nuclear charge. The upper positronic eigenvalue is plotted as a function of total charge in Fig.2.5. We observe a perfect linear fit, which confirms the physical picture outline above. Finally, I have considered the behaviour of the "unphysical" solutions in the presence of an attractive potential. I performed a series of Dirac-Fock calculations on the neon atom, but with variable electron occupation, so as to vary the total charge of the system. The upper positronic eigenvalue is plotted as a function of total charge in Fig.2.6. In the part of the plot corresponding to positive total charge we see the same linear fit as before. For negative charges, however, a non-linear deviation is observed. It corresponds to weakly bound positrons with energies larger than $-2mc^2$.

We can conclude that the "unphysical" positron solutions have a very physical be-

---

[3] This has the well-known consequence that the lower virtual electronic Dirac-Fock orbitals are rather diffuse and not particularly well suited for correlation[156, 157]. A correlated method based on a truncated one-particle basis should therefore be performed using natural orbitals [158] or modified virtual orbitals [159].

haviour in terms of delocalized positrons. We also see that the appearance of positronic solutions with energies above $-2mc^2$ in calculations on systems with total negative charge, does not necessarily indicate a basis set failure, but is a consequence of the positive charge of the positron.

The "unphysical" solutions, by their nature, do not perturb the electronic solutions, but they reflect a redundancy in the small component basis. They can be deleted from the molecular orbital space by the following simple method:

The Dirac-Fock equations are solved iteratively. In each SCF iteration we solve the general eigenvalue problem

$$\mathbf{F}\mathbf{c} = \varepsilon\mathbf{S}\mathbf{c} \tag{2.42}$$

The overlap matrix $\mathbf{S}$ appears since we are working within the non-orthogonal atomic orbital (AO) basis. The general eigenvalue problem is solved by first transforming to orthonormal (MO) basis. We may do so by a canonical orthonormalization $\mathbf{V}$ [160] constructed from eigenvalues s_i and eigenvectors O_{*i} of the overlap matrix

$$\mathbf{V} = \mathbf{O}\mathbf{s}^{-1/2}; \qquad \mathbf{O}^T\mathbf{S}\mathbf{O} = \mathbf{s}, \quad s_{ij} = s_i\delta_{ij} \tag{2.43}$$

The eigenvalue problem is then reduced to a standard eigenvalue problem

$$\mathbf{F}'\mathbf{c}' = \varepsilon\mathbf{c}'; \qquad \mathbf{F}' = \mathbf{V}^\dagger\mathbf{F}\mathbf{V}, \mathbf{c}' = \mathbf{V}^{-1}\mathbf{c} \tag{2.44}$$

solved by a complex or quaternion diagonalization, depending in the formalism used. The MO-coefficients are recovered by the backtransformation

$$\mathbf{c} = \mathbf{V}\mathbf{c}' \tag{2.45}$$

In the transformation to MO-basis we could of course have used any transformation that orthonormalizes the basis, for example MO-coefficients from a given SCF iteration. The advantage of the canonical orthonormalization is that it allows a straightforward deletion of numerical dependencies introduced by large basis expansions. We simply delete columns of the MO-transformation matrix $\mathbf{V}$ Eq.(2.43) corresponding to eigenvalues of the overlap matrix below a selected treshold. From this we see that the MO-basis need not be of the same size as the AO-basis. We can use the MO-transformation to project the AO-basis down onto a MO-basis spanning a smaller space. This reduces the number of variational parameters and might thereby improve convergence.

The deletion of "unphysical" solutions can be embedded in the MO-transformation. We first solve the free-particle Dirac equation in the current basis. This amounts to solving an eigenvalue problem of the form Eq.(2.42), where $\mathbf{F}$ is the matrix representation of the free-particle Hamiltonian in AO-basis. At this point we can use the canonical orthonormalization to remove linear dependencies from the MO-space. The "unphysical" positronic solutions are readily identified by their eigenvalues $-2mc^2$ and are discarded. The remaining set of MO-coefficients is then used as MO-transformation matrix in all subsequent SCF iterations, thereby ensuring the removal of both linear dependencies and "unphysical" solutions. The method shows that projection operators are straightforwardly emebedded in the MO-transformation. We will exploit this in the next section.

2.2.6 Variational stability of the Dirac-Coulomb Hamiltonian

We now return to the controversy surrounding the validity of the Dirac-Coulomb Hamiltonian. The discussion was initiated by Brown and Ravenhall [161]. They considered a system of two non-interacting bound electrons. If the electron-electron interaction is turned on as a perturbation, there will be an infinite number of degenerate states consisting of one electron from the positive continuum and one positron from the negative continuum. Thus the original bound electronic configuration will evolve into a completely delocalized two-particle state. This is referred to as the "continuum dissolution" or the "Brown-Ravenhall disease". The solution proposed by Brown and Ravenhall was to restrict the Hamiltonian to positive energy states by surrounding it with projection operators and thereby neglect all pair creation processes. This solution was further expounded by Sucher [133, 134, 135, 28]. There are, however, several possible choices of projection operators. They may be defined in terms of positive energy solutions to the free particle Dirac equation; this corresponds to the 'free' picture of QED. Alternatively, one can define the projection operators in terms of positive energy solutions of the external field Dirac equation (the "Furry" picture [136]). In chemical applications, the external field would be the molecular field defined by the nuclei. A third solution was proposed by Mittleman [162], namely to construct the projection operator iteratively from the positive energy solutions of the Dirac-Fock equation. This has been denoted the "fuzzy" picture [28]. These proposals have been reviewed by Kutzelnigg [163].

The controversy and the methodological difficulties surrounding the Dirac-Coulomb Hamiltonian are largely resolved today, as manifested by the routine application of the Dirac-Coulomb Hamiltonian in relativistic molecular calculations. We first note that the no-pair approximation can only be made with reference to some 1-particle basis, which defines what is electrons and what is positrons. The variational methods of relativistic molecular calculations at some stage involve the construction of a one-particle basis by the solution of an effective one-electron Dirac equation, albeit with a non-local potential. Electronic solutions appear as excited states in the spectrum of the effective one-electron operator. As pointed out by Talman [164, 165], they are found by application of a minimax principle, where the energy is minimized with respect to rotations into the virtual electron space, and maximized with respect to rotations into the the positron space. With a balanced basis, the electronic and positronic solutions are well separated and readily identifiable. Once electrons and positrons have been defined, configurational expansions can be limited to electronic configurations, in accordance with the no-pair approximation.

In the Dirac-Fock method, the electronic solutions are straightforwardly obtained by vector selection. This corresponds to the use of projection operators within the "fuzzy" picture. The projection operators need not, however, be defined explicitly. The performance of projection operators defined within the "free" and "Furry" picture can be tested within the Dirac-Fock model using the method for embedding projection operators in the MO-transformation, outlined in the previous section. "Free" projection operators are obtained by first solving the free-particle Dirac equation in the AO-basis. We then discard the positronic solutions and use the remaining vectors set as our MO-transformation matrix in all subsequent SCF iterations. The "Furry" projection operators are obtained in analogous manner, but starting from the molecular field Dirac equation. The performance of various projection operators has been studied by calculations on the radon atom ($Z = 86$) in an uncontracted Cartesian Gaussian dual family basis. The large component basis (22s19p14d9f) consisted of 253 functions, and the small component basis (19s22p19d14f9g), generated by the unrestricted kinetic balance condition, consisted of 474 functions. The results are presented in Tab.2.2.6. The "free" projection operator is seen to perform poorly, which is not unexpected [166]. The introduction of an external field leads to a "dressing" of the free electrons and positrons. The localized bound electrons are of a quite different nature than the delocalized free electrons. The "Furry" projection operator, however, is seen to perform rather well. We can understand this by the fact that the small component density is localized mostly to the nuclear region, subject to the almost unscreened nuclear charge. The introduction of the electron-electron interaction therefore does not drastically change the small component density. In Tab.2.2.6 we have also included result obtained by the method described in the previous section, where a one-to-one matching of the large and small component basis is obtained by projecting out the "unphysical" free positron solutions. Thes results are seen to be equivalent to the results obtained by the basis set extension of the atomic code GRASP [117, 96], where a one-to-one matching of large and small component basis functions is obtained by restricted kinetic balance defined in terms of Gaussian 2-spinors (see paper I).

2.2.7 Direct SCF: a numerical example

In paper I we have presented the theory of the 4-component SCF method. A key ingredient in direct SCF is the screening of contributions to the Fock matrix to reduce the time spent in each SCF iteration. In this section we illustrate the method by a numerical example. Let us first briefly review the screening implemented in DIRAC.

Integrals are generated in batches (K,L,M,N) defined by shell indices and screening is based on a threshold τ. In each SCF iteration we generate a density matrix over shell indices to be kept in memory

$$D_{KL} = \max\left(D_{\kappa\lambda;\Lambda}\right); \qquad \kappa \in K, \lambda \in L, \Lambda \in [0,3] \tag{2.46}$$

Table 2.3: Total (E_{DF}) and orbital energies (in Hartrees) of the radon atom from Dirac-Fock calculations using various schemes. UKB = the large and small component bases related by unrestricted kinetic balance. RKB = "unphysical" free positronic solutions projected out of the MO-basis. VEXT = all positronic solutions of the molecular field Dirac equation projected out of the MO-basis. FREE = all positronic solutions of the free particle Dirac equation projected out of the MO-basis. GRASP = results from a calculation performed with the basis set extension of GRASP, which employs restricted kinetic balance defined in terms of a Gaussian 2-spinor basis.

	UKB	RKB	$VEXT$	$FREE$	$GRASP$
E_{DF}	−23601.879021	−23601.929261	−23601.911484	−24387.589625	−23601.930828
$1s_{1/2}$	−3641.175215	−3641.172817	−3641.189420	−3904.888545	−3641.173
$2s_{1/2}$	−668.806797	−668.805891	−668.810237	−707.522900	−668.8059
$2p_{1/2}$	−642.279195	−642.295351	−642.284778	−671.753287	−642.2954
$2p_{3/2}$	−541.101773	−541.100868	−541.105285	−553.106216	−541.1008
$3s_{1/2}$	−166.829782	−166.829514	−166.830532	−176.754326	−166.8295
$3p_{1/2}$	−154.879012	−154.883240	−154.880136	−162.755672	−154.8832
$3p_{3/2}$	−131.728395	−131.728144	−131.729117	−135.388427	−131.7281
$3d_{3/2}$	−112.562535	−112.563187	−112.563135	−116.264914	−112.5632
$3d_{5/2}$	−107.756490	−107.756249	−107.756985	−110.473356	−107.7562
$4s_{1/2}$	−41.310404	−41.310334	−41.310574	−44.266294	−41.31032
$4p_{1/2}$	−36.012965	−36.014083	−36.013217	−38.359034	−36.01407
$4p_{3/2}$	−30.117674	−30.117608	−30.117829	−31.339004	−30.11759
$4d_{3/2}$	−21.544561	−21.544774	−21.544668	−22.671517	−21.54476
$4d_{5/2}$	−20.435777	−20.435722	−20.435860	−21.317706	−20.43571
$4f_{5/2}$	−9.190776	−9.190751	−9.190788	−9.911893	−9.190735
$4f_{7/2}$	−8.925160	−8.925127	−8.925167	−9.493619	−8.925110
$5s_{1/2}$	−8.405941	−8.405924	−8.405975	−9.317665	−8.405917
$5p_{1/2}$	−6.405063	−6.405298	−6.405110	−7.128608	−6.405292
$5p_{3/2}$	−5.172874	−5.172857	−5.172902	−5.636034	−5.172849
$5d_{3/2}$	−2.186505	−2.186538	−2.186518	−2.566585	−2.013540
$5d_{5/2}$	−2.013557	−2.013546	−2.013567	−2.348793	−2.186532
$6s_{1/2}$	−1.068464	−1.068461	−1.068470	−1.393778	−1.068460
$6p_{1/2}$	−0.536667	−0.536697	−0.536673	−.795442	−0.536698
$6p_{3/2}$	−0.381745	−0.381742	−0.381749	−.590526	−0.381740

and for each integral batch we define

$$D_{max} = \max\left(D^C_{max}, D^E_{max}\right); \qquad \left\{ \begin{array}{lcl} D^C_{max} & = & 4 \cdot \max\left(D_{KL}, D_{MN}\right) \\ D^E_{max} & = & \max\left(D_{NL}, D_{ML}, D_{NK}, D_{MK}\right) \end{array} \right. \tag{2.47}$$

We furthermore calculate a strict upper limit to the size of two-electron integrals in the batch by

$$I_{KLMN} = \max\left[G_{\kappa\lambda} G_{\mu\nu}\right]; \qquad G_{\kappa\lambda} = \sqrt{(\kappa\lambda \mid \kappa\lambda)}; \quad \kappa \in K, \lambda \in L, \mu \in M, \nu \in N \tag{2.48}$$

In DIRAC screening is performed in two steps:

1. *Prescreening of integral batches*
   We only calculate batches for which
   $$D_{max} I_{KLMN} \geq \tau \tag{2.49}$$

2. *Prescreening of contributions*
   If the batch is calculated, we perform a second separate screening on Coulomb
   $$D^C_{max} I_{KLMN} \geq \tau \tag{2.50}$$
   and exchange integrals
   $$D^E_{max} I_{KLMN} \geq \tau \tag{2.51}$$
   before feeding the integrals to the routine for the construction of the Fock matrix. If the exchange contributions are screened out, we calculate only the Coulomb contributions to the Fock matrix, and vice versa.

We illustrate the method by an example. I have performed a 4-component direct Dirac-Fock calculation on the diatomic interhalogen BrI at the experimental gas phase bond length 248.5 pm [167] in a basis of uncontracted Cartesian Gaussian dual family basis. The basis sets are given in Tab.2.2.7. The Dirac-Fock calculation was performed in the following manner: In the initial SCF iterations the two-electron Fock matrix is constructed from only LL integrals. SL integrals are not introduced into the SCF process until the convergence on total energy is less than $1.0 \cdot 10^{-4}$ Hartrees. Likewise, SS integrals are not introduced until the convergence on total energy is less than $1.0 \cdot 10^{-6}$ Hartrees. Screening is based on the differential density approach with threshold $\tau = 1.0 \cdot 10^{-8}$. However, in a SCF iteration where a new integral class is introduced, differential densities are replaced by true densities, in order to avoid the introduction of errors, as discussed in paper I.

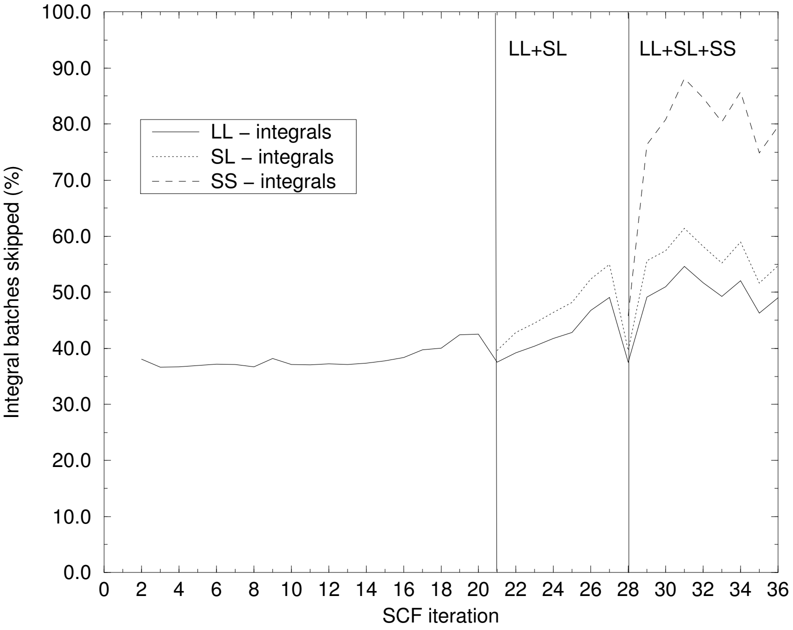


Figure 2.7: 4-component direct SCF calculation on BrI: Percentage of integral batches screened out in the first step.

Table 2.4: Basis sets of uncontracted Cartesian Gaussian used in the calculation on BrI. χ^L and χ^S refer to the large and small components basis sets respectively. The total number N_T of basis functions is the sum of the number of large (N_L) and small (N_S) basis set functions.

	χ^L	N_L	χ^S	N_S	N_T
Br	16s14p9d	112	14s16p14d9f	236	348
I	20s17p11d	249	17s20p17d11f	289	426

Table 2.5: Effects of screening in 4-component direct SCF calculation on BrI. The CPU-time refers to the average (over SCF iterations) absolute time(h:min:s) and the time relative to the unscreened case for processing integrals and Fock matrix.

	Integral batches screened out (average)			Average CPU-time	
	First step	Second step			
		Coulomb	exchange		
LL integrals	41.8%	1.6%	1.3%	0:09:10	90.5%
SL integrals	51.4%	0.5%	10.9%	1:00:25	67.2%
SS integrals	77.4%	1.5%	8.2%	0:58:27	29.7%

The effects of screening on the iterative process are summarized in Tab.2.5 and visualized in Figs. 2.7,2.8 and 2.9. On average 77.4% of the SS-integral batches are screened out, so that the time spent on generating the integrals and processing their contributions to the Fock matrix is reduced by a factor of more than three, compared to the corresponding calculation without screening. On average, more time is in fact spent on processing the SL-integrals. The effect of screening Coulomb contributions is small, but this is reasonable due to the long-range behavior of the Coulomb interaction. The screening of exchange contributions , on the other hand, has an appreciable effect. The effect is slightly larger for SL integrals than for SS integrals, but this is probably due to the fact that so many batches of SS integrals are screened out in the first step. The "dip" in the curves at iteration 21 and 28 in Fig2.7 gives some indication of the effect the differential density approach, because in these two iterations new integral classes are introduced and the differential densites replaced by absolute densities. In iteration 35, near convergence, there is a reduction in the number of integral batches that are screened out, but this is due to the use of a dynamic threshold τ; when the energy convergence passes below the threshold τ, the threshold is adjusted down accordingly to increase the precision in the Fock matrix construction.

Further studies on the effect of screening in 4-component direct SCF are needed to optimize the process. Also, the routine for the construction of Fock matrices is currently somewhat too slow.

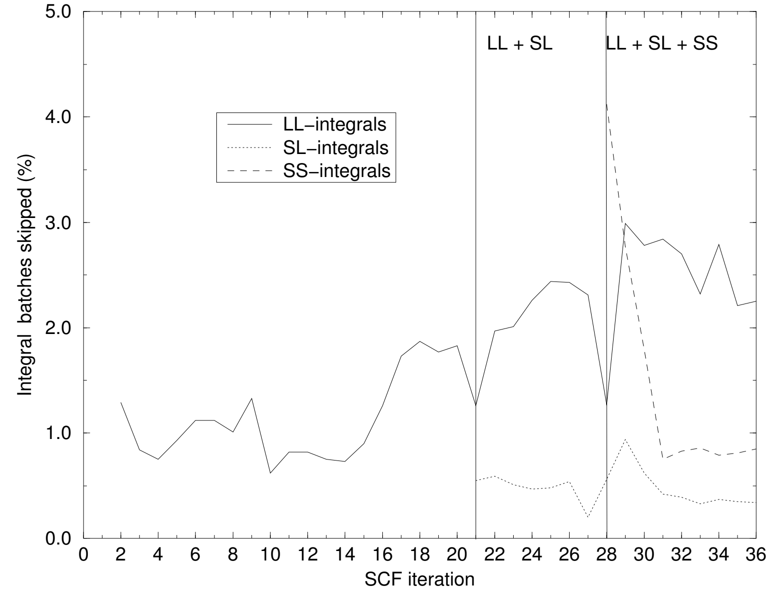


Figure 2.8: 4-component direct SCF calculation on BrI: Percentage of integral batches for which Coulomb contributions were screened out in second step

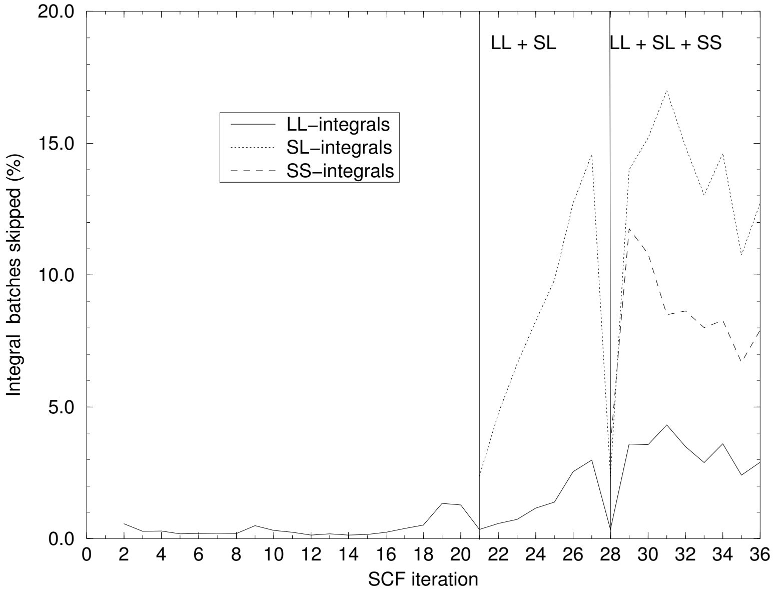


Figure 2.9: 4-component direct SCF calculation on BrI: Percentage of integral batches for which exchange contributions were screened out in second step

2.3 Concluding remarks

> *. . . after a Dirac lecture, the session chairman asked the audience if they had any questions. A person stood up and said 'I did not understand the derivation of . . . etc, etc.' Paul made no response, the chairman asked, 'aren't you going to answer the question?' Paul said, 'that was a statement, not a question.'*
>
> J.E.LANNUTTI *(1987) [2]*

In this thesis contributions to the theory of 4-component relativistic molecular calculations have been presented. The theory of the 4-component direct SCF method has been implemented in the code DIRAC. The code is currently being applied to several molecular calculations. We have also presented the theory of multi-configurational self-consistent field (MCSCF) method. As discussed in previous sections, the need multi-configurational approaches is amplified in the relativistic domain due to the fine structure provided by the spin-orbit interaction. The 4-component MCSCF method is under implementation in cooperation with K.G. Dyall and H.J.Aa.Jørgensen.

The 4-component methods provide a relativistic description of molecules. The operators have a simple structure, and are well-defined for a wide range of molecular properties. Work is therefore in progress in the development of 4-component methods for first and second order molecular properties. For properties dependent on the electron density in the nuclear region, 4-component methods will be relevant even for fairly light systems.

The main difficulty of the 4-component methods lies in their computational expense. The situation has been improved with the introduction of direct SCF methods, which makes it possible to perform relativistic molecular calculations on work stations. There is, however, a need to reduce the large number of integrals stemming from the small component basis. The prospects for integral approximations looks good, due to the localized atomic nature of the small component density. Progress in this area would open up a wide range of chemical problems for study by 4-component methods. For large systems such investigations will benefit from the explicit calculations of molecular gradients and the use of second order optimization methods. This should therefore be research areas of high priority.

We may conclude that the future for 4-component methods looks bright.

Part II

Papers

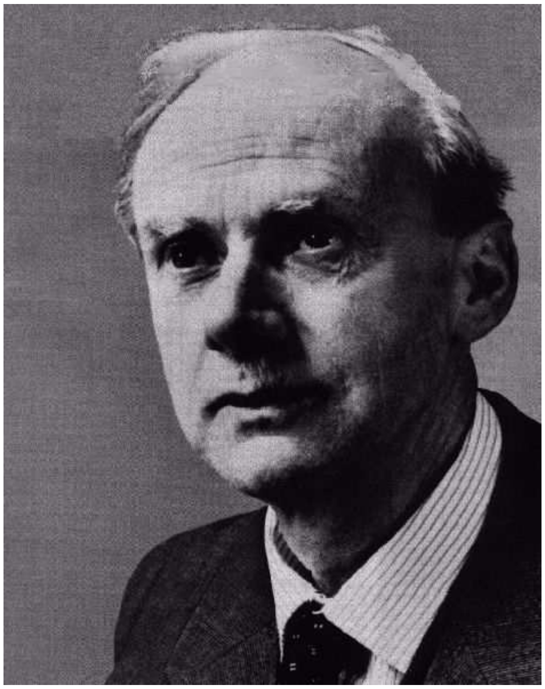
Part III

Appendices

Appendix A

DIRAC : documentation

P	**rogram**
•	**for**
A	**tomic**
•	**and**
M	**olecular**
D	**irect**
I	**terative**
R	**elativistic**
A	**ll-electron**
C	**alculations**



A.1 General overview

Version 2.0. Last update: Nov 14 1995 - tsaue@kelvin.uio.no

DIRAC is a FORTRAN code for relativistic molecular calculations based on the Dirac-Coulomb (-Gaunt) Hamiltonian. It solves the the 4-component Dirac-Fock(DF) equations by the Self Consistent Field (SCF) iterative procedure and provides tools for analysis of the converged wave function. **DIRAC** is constructed around **HERMIT** , a highly efficient code for generating one- and two-electron integrals over a basis of Cartesian Gaussian functions. The two-electron integrals naturally split into three classes:

- LL-integrals — $(LL\,|\,LL)$
- SL-integrals — $(SS\,|\,LL)$
- SS-integrals — $(SS\,|\,SS)$

(possibly extended by Gaunt integrals). For each integral type two integral processing modes are accessible to the user

- *conventional mode*: symmetry-adapted integrals are stored on disk
- *direct mode*: integrals are regenerated whenever they are needed

In the latter mode a differential density matrix approach may be used to reduce the number of integrals calculated in each iteration. This number may be reduced further by including SL- and/or SS-integrals only at an advanced stage in the SCF-iterations, determined by convergence criteria or by specifying the iteration at which to include the integrals. Convergence may be enhanced by damping the Fock matrix or by Direct Inversion of Iterative Subspaces (DIIS).

In the present version only the large component basis needs to be defined, the small components then being generated by the kinetic balance prescription. Restricted kinetic balance may be enforced by deleting unphysical solutions in the positron spectrum of free electron solutions. The nuclear chrage distribution is represented by a Gaussian function to avoid the sinuglarities introduced by point nuclei. Time reversal symmetry is implemented using quaternion algebra, whereas spatial symmetry is restricted to the binary groups, that is D_{2h} and subgroups.

DIRAC *is an experimental code, subject to continous change.*

A.2 Recent modifications

Nov 11 1995	Overlap selection implemented with keyword `OVLSEL`.
Nov 11 1995	Keyword `NOSMLV` turns off the small component nuclear attraction integrals and thereby the spin-orbit interaction in the field of the nuclei.
Nov 3 1995	File `DFCYCL` containing SCF history is formatted.

Nov 3 1995	Convergence acceleration: By default DIIS is activated using a very large threshold (`DIISTH`). When DIIS can not be used damping is activated using a default damping factor 0.25. DIIS can be turned off using the keyword `.NODIIS`. The DIIS threshold can be modified using the keyword `.DIISTH`. Damping can be turned off using the keyword `.NODAMP`. The damping factor can be modified using the keyword typ .DAMPFC.
Nov 3 1995	Restart facilities have been made more robust, see section A.7.
Oct 31 1995	Even when an integral batch is generated it is possible to have separate screening of Coulomb and exchange contributions. Exchange contributions are generally more local and easier to screen out. This feature is activated by the keyword `.CEDIFF`.
Oct 31 1995	All free positronic solutions may be projected out of the MO-space using the keyword `.FREEPJ`
Oct 31 1995	All positronic solutions of the one-electron Fock matrix may be projected out of the MO-space using the keyword `.VEXTPJ`

A.3 Installing the program

The program can be installed on a number of different types of computers. It is present as a series of master files that are processed by the `UPDATE` code to handle machine-specific features. The master files consists of three categories

`dir*.u`	**DIRAC** - files
`*her*.u`	**HERMIT** - files
`gp*.u`	library - files

In addition there are `.cdk` - files containing machine-specific features and COMMON-blocks:

`gen.cdk`	- general features
`dirac.cdk`	- features pertaining specifically to **DIRAC**
`aba.cdk`	- features pertaining specifically to **HERMIT**

`gpg08.cdk`	- features pertaining specifically to library routines

Two shell scripts are available for installing the program:

`lag`	- install one or more main modules: **HERMIT** , **DIRAC** or library files
`add`	- add a specific master file

A.4 Running the program

Two input files are needed to run **DIRAC** :

- The **basis file** defines the basis set, nuclear configuration and symmetry.
- The **menu file** defines the calculation.

At the start of any calculation the basis file is processed and then various modules are activated based on the information given in the menu file. One may therefore run a sequence of calculations based on the same basis file. Intermediate files are kept to a minimum. For instance, one may first run a DF-calculation which gives a set of MO-coefficients defining the converged wave function. Population analysis may be performed in a separate calculation and then requires only the file of coefficients in addition to the basis file.

The basic UNIX command for running the program is

```
dirac.x < {menufile} > {outputfile}
```

The basis file must be present as `MOLECULE.INP`.

The various program modules have different memory requirements. Memory may be allocated dynamically by setting an environmental variable `DIRWRK` which defines the number of 8-byte words needed in the calculation, e.g.

```
  setenv DIRWRK 30000000
```

means that 30 Mw of memory are to be allocated[1] Default memory is specified by the variable `LWORK` which may be set in the `dirac.cdk` - file.

A C-shell script `dhf` is available for automatization of calculations:

```
Usage: dhf [flags] file
   Flags:
     -incmo     : Copy DFCOEF to   work area
     -utcmo     : Save DFCOEF from work area
```

---

[1] 1 Mw = 8,000,000 bytes = 7.63 MB. 1MB = $(1024)^2$ bytes.

```
-infck     : Copy DFFCK2 to   work area
-utfck     : Save DFFCK2 from work area
-mw mem    : set memory(in megawords)
-rmwrk     : remove work area after calculation
-run name  : File suffix for output
```

A.5 Input files

Common to both the menu file and the basis file is that any line that begins with the hash symbol `#` is interpreted as a comment. This feature may be used for the insertion of comments and in the menu file to "turn off" keywords.

A.5.1 Menufile

The menu file defines the calculation and has the general structure

```
*DIRAC
  < keywords >
< chapter >
   < keywords >
< chapter >
   < keywords >
 ...
*END OF
```

Each `chapter` has an asterisk (`*`) in the initial position and generally refer to a program module. For each `chapter` a set of `keywords` may be specified, possibly with additional arguments. The set of `chapters` and `keywords` allows the user great flexibility in defining the current calculation.

1. ***DIRAC**

   (a) **Job assignment**
   In this section the program modules to be called are defined. It is recommended to start a new set of calculations with all modules turned off in order to check the basis file processing.

`.TITLE`	title line *Arguments:* Title line (max. 50 characters) *Default:* `DIRAC: No title specified !!!`
`.INPTES`	input test: no job modules called *Default:* `INPTES = .FALSE.`
`.DHFCAL`	perform Dirac-Fock calculation *Default:* `DODHF = .FALSE.`
`.DHFANA`	analyze Dirac-Fock wave function *Default:* `DOANA = .FALSE.`

   (b) **Job control**

`.DIRECT`	- direct evaluation of two-electron integrals. *Arguments:* Integers `ILL,ISL,ISS` `IXX = 1`(on)/`0`(off) (`XX = LL,SL` or `SS`) *Default:* `ILL = ISL = ISS = 0`
`.ONESYS`	ignore two-electron part *Default:* `ONESYS = .FALSE.`
`.NSYM`	number of fermion ircops (one or two). *Default:* `NSYM = 1`
`.URKBAL`	unrestricted kinetic balance *Default:* `URKBAL = .FALSE.` Restricted kinetic balance. This is imposed by deleting unphysical solutions from the free particle positronic spectrum.
`.FREEPJ`	project out all free positronic solutions from the MO-space *Default:* `.FALSE.`
`.VEXTPJ`	project out all external field positronic solutions from the MO-space *Default:* `.FALSE.`
`.SPHTRA`	transformation to spherical harmonics embedded in transformation to orthonormal basis; totally symmetric contributions deleted. *Arguments:* Integers `ISPHL,ISPHS` — `ISHP = 1(on)/0(off)` *Default:* `ISPHL = 1,ISPHS = 0`. Note that with `ISPHL = 1` and *restricted* kinetic balance, the correct transformation of the small components is automatically imposed.
`.CVALUE`	reset the value of light *Arguments:* `CVAL` *Default:* `CVAL = 137.03604D0`
`.PTNUC`	use point nucleus *Default:* The nuclear charge distribution is represented by a Gaussian function.
`.TIMINT`	time integral evaluation *Default:* `TIMINT = .FALSE.`
`.OVLTOL`	thresholds for linear dependence in large and small components *Arguments:* Real `STOL(1),STOL(2)` *Default:* Large: `STOL(1) = 1.0D-6`. Small: `STOL(2) = 1.0D-8`

(c) Print levels

`.PRINT`	general print level *Arguments:* Integer `IPRGEN`

	Default: `IPRGEN = 0`
`.IPRONE`	print level for one-electron integrals *Arguments:* Integer `IPRONE` *Default:* `IPRONE = 0`
`.IPRTWO`	print level for two-electron integrals *Arguments:* Integer `IPRTWO` *Default:* `IPRTWO = 0`

2. ***READIN** — Input module : Read and process menu and basis files.

`.PRINT`	general print level in input module *Arguments:* Integer `IPREAD` *Default:* `IPREAD = 0`
`.MAXPRI`	maximum number of primitive functions in a given block in basis file *Arguments:* Integer `MAXPRI` *Default:* `MAXPRI = 15`

3. ***DHFCAL** — Perform Dirac-Fock calculation.

(a) **DHF – occupation**

`.NELECT`	for each fermion ircop, give number of electrons *Arguments:* Integers `(NELEC(I),I=1,NSYM)` *Default:* `NELEC(1) = NELEC(2) = 0`

(b) **Print levels**

`.PRINT`	general print level *Arguments:* Integer `IPRDHF` *Default:* `IPRDHF = 0`

(c) **Trial function**

A DF-calculation may be initiated in three different ways:

- using **MO-coefficients** from a previous calculation.
- using **two-electron Fock matrix** from a previous calculation; this may be thought of as starting from a converged DHF potential
- using coefficients obtained by diagonalization of the one-electron Fock matrix: the **bare nucleus approach**.

Default is to start from MO-coefficients if the file `DFCOEF` is present. Otherwise the bare nucleus approach is followed. In all three cases linear dependencies are removed in the 0th iteration.

`.TRIVEC`	start SCF-iteratons from vector file

`.TRIFCK`	start SCF-iterations from two-electron Fock matrix from previous calculation

(d) **Convergence criteria**
Three different criteria for convergence may be chosen:

- the difference in total energy between two consecutive iterations
- the largest absolute difference in the total Fock matrix between two consecutive iterations
- the largest element of the DIIS error vector $\mathbf{e} = [\mathbf{F}, \mathbf{D}]$ (in MO-basis).

The change in total energy is approximately the square of the largest element in the error vector or the largest change in the Fock matrix. *Default* is convergence on error vector with threshold `SCFCNV = 1.0D-6` . Alternatively, the iterations will stop at the maximum number of iterations.

`.MAXITR`	maximum number of SCF - iterations *Arguments:* Integer `MAXITR` *Default*: `MAXITR = 50`
`.ERGCNV`	threshold for convergence on total energy *Arguments:* Real `SCFCNV`
`.EVCCNV`	converge on error vector *Arguments:* Real `SCFCNV`
`.FCKCNV`	converge on largest absolute change in Fock matrix *Arguments:* Real `SCFCNV`

(e) **Convergence acceleration**
It is imperative to keep the number of SCF-iterations at a minimum. This may be achieved by convergence acceleration schemes.

- **Damping** The simplest scheme is damping of the Fock matrix that may remove oscillations. In iteration `n + 1` the Fock matrix to be diagonalized is:

$$\mathbf{F}' = (1-c)\mathbf{F}_{n+1} + c\mathbf{F}_n; \quad c --- \text{damping factor} \tag{A.1}$$

- **DIIS** (Direct Inversion of iterative Subspaces) may be thought of as generalized damping involving Fock matrices from many iterations. Damping factors are obtained by solving a simple matrix equation involving the B-matrix constructed from error vectors (approximate gradients).

In **DIRAC** DIIS takes precedence over damping.

`.DIISTH`	change default threshold for initiation of DIIS, based on largest element of error vector

Arguments: Real `DIISTH` — convergence threshold for initiation of DIIS
Default: A very large number.

`.MXDIIS`	maximum dimension of B-matrix in DIIS module *Arguments:* Real `MXDIIS` — maximum dimension of B-matrix *Default:* `MXDIIS = 15`
`.NODIIS`	do NOT perform Direct Inversion of Iterative Subspaces (DIIS) *Default:* DIIS is activated.
`.DAMPFC`	change default damping factor *Arguments:* Real `DAMPFC` — damping factor *Default:* `DAMPFC = 0.25`.
`.NODAMP`	do NOT perform damping of Fock matrix *Default:* Damping is activated, but DIIS takes precedence.

(f) **State selection**

Convergence can be improved by selection of vectors based on overlap with vectors from a previous iteration. This method mayu also be used for convergence to some excited state. If **DIRAC** starts on a vector set , this vector set forms the criterium for overlap selection, otherwise the criterium are the vectors from the first iteration. Vector selection based on vectors generated by the bare nucleus approach are not recommended.

`.OVLSEL`	activate overlap selection *Default:* No overlap selection.

(g) **Iteration speedup**

The total run time may be reduced significantly by reducing the number of integrals to be processed in each iteration:

- **Screening on integrals:** Thresholds may be set to eliminate integrals below the threshold value. The threshold for LL-integrals is set in the basis file, but this threshold may be adjusted for SL- and SS-integrals by threshold factors:
  - Threshold for LL-integrals: `THRS`
  - Threshold for SL-integrals: `THRS*THRFAC(1)`
  - Threshold for SS-integrals: `THRS*THRFAC(2)`
- **Screening on density:** In direct mode further reductions are obtained by screening on the density matrix as well. This becomes even more effective if one employs **differential densities**, that is

$$\Delta\mathbf{D} = \mathbf{D}_{n+1} - \mathbf{D}_n \tag{A.2}$$

- **Neglect of integrals:** The number of integrals to be processes may be reduced even further by adding SL- and SS-integrals only at an advanced stage in the DHF-iterations, as determined either by the number of iterations or by energy convergence. The latter takes precedence over the former.

`.THRFAC`	adjust integral thresholds for SL- and SS-integrals. *Arguments:* Reals `THRFAC(1),THRFAC(2)` *Default:* `THRFAC(1) = THRFAC(2) = 1`
`.DNSTHR`	threshold for screening on density matrix *Arguments:* `DNSTHR` *Default:* Real `DNSTHR = 1.0D-10`
`.CEDIFF`	separate density screening of Coulomb and exchange contributions *Default:* `.FALSE.`
`.NODSCF`	do not perform SCF - iterations with differential density matrix *Default:* use differential density matrix in direct SCF.
`.CNVINT`	set threshold for convergence before adding SL- and SS-integrals to SCF-iterations. *Arguments:* Reals `CNVINT(1)`(SL),`CNVINT(2)`(SS) *Default:* Very large numbers.
`.ITRINT`	set number of iterations before adding SL- and SS-integrals to SCF-iterations. *Arguments:* Integers `ITRINT(1)` (SL),`ITRINT(2)`(SS) *Default:* `ITRINT(1) = ITRINT(2) = 1`
`.NOSMLV`	turn off small component nuclear attraction integrals; this turns off the spin-orbit interaction from the field of nuclei. If SL two-electron integrals are turned off as well, all spin-orbit interaction is cancelled. *Default:* Integrals are included.

(h) **Output control**

`.VECPRI`	separate control of printing of large and small components *Arguments:* Integers `IPRVEC(1)`(large),`IPRVEC(2)`(small). *Default:* No vectors printed.
`.EIGPRI`	control printing of electron and positron solutions *Arguments:* Integers `IPREIG(1)` (electron),`IPREIG(2)`(positron) *Default:* Electronic eigenvalues printed.
`.SPINOR`	for each fermion ircop, give number of spinors to print. *Default:* the occupied electronic solutions. *Arguments:* Integers `(NSPI(I),I=1,NSYM)`

`.OFFSET` for each fermion ircop, offset for start address for vectors to print.
Arguments: Integers `(ISPI(I),I=1,NSYM)`
Default: 0 , meaning the first electron solution. (Positron solutions may be printed by giving negative offsets).

4. ***DHFANA** — Analyze Dirac-Fock wave function. Mulliken population analysis is performed in AO-basis. The analysis is based on the concept of *labels*. Each basis function is labeled by its functional type and center. The labels are given in output. A set of primitive labels may be collected to group labels as specified by the user.

`.PRINT` general print level
Arguments: Integer `IPRANA`
Default: `IPRANA = 0`

`.MULPOP` give Mulliken gross populations
Default: `DOMULP = .FALSE.`

`.NETPOP` give Mulliken gross and net/overlap populations
Default: `DONETP = .FALSE.`

`.LABDEF` defined labels for use in Mulliken population analysis
Arguments: Integer `NCLAB` - number of labels to define

```
DO I = 1,NCLAB
  READ(LUCMD,'(A12,I5)') CLABEL(I),NGRPS
  READ(LUCMD,*) (IBUF(J),J=1,NGRPS)
  DO J = 1,NGRPS
    ICLAB(IBUF(J)) = I
  ENDDO
ENDDO
```

`.ADDSML` use default labels for small components
Default is to gather all small component functions belonging to a given center.

`.SPINOR` for each fermion ircop, give number of spinors to analyze.
Default: the occupied electronic solutions.
Arguments: Integers `(NSPI(I),I=1,NSYM)`

`.OFFSET` for each fermion ircop, offset for start address for vectors to analyze.
Arguments: Integers `(ISPI(I),I=1,NSYM)`
*Default:*0 , meaning the first electron solution. (Positron solutions may be analyzed by giving negative offsets).

A.5.2 Basisfile

The basis file defines the present basis set, molecular geometry and the symmetry of the system. A summary of the basis file structure is given in Tab. A.5.2.

Card 1 : `INTGRL` — a keyword that identifies the file.

Card 2 : `TITLE(1),TITLE(2) (A72/A72)` Two title lines.

Card 3 : `CRT,NONTYP,NSYMOP,(SYMOP(I),I=1,3),THRS`
`(BN,A1,I4,I5,3A3,1X,D10.2)`

`CRT` - must be set to `'C'` to indicate that Cartesian Gaussians are to be used.

`NONTYP` - specify the number of atomic types

`NSYMOP` - specify the number of generators of the symmetry group

`SYMOP(I)` - Symmetry is restricted to restricted to the binary groups, that is D_{2h} and subgroups, which means that a symmetry operation acting on the main axes (x,y,z) will at most reverse their direction. A group generator is therefore identified by a 3-character string that specifies the axes reversed under its operation. Examples are given for the eight binary groups in Tab. A.5.2.

`THRS (D10.2)` Threshold for LL-integrals. Separate thresholds for SL- and SS-integrals may be specified in the menu file.

- *For each atomic type:*

  Card 3.1 `Q,NONT(I),QEXP (BN,F10.0,I5,F20.5)`

  `Q` - nuclear charge

  `NONT(I)` - number of symmetry independent centers

  `QEXP` Gaussian exponent for nuclear charge distribution (if zero, then default is used).

  - *For each symmetry independent center:*

    Card 3.1.1 `NAMN(NUCIND),[CORD(J,NUCIND), J = 1,3] (BN,A4,3F20.0)`

    `NAMN` - name of nuclear center

    `CORD(J)` - x-,y-and z-coordinate of nuclear center

  Card 3.2 `BSET,IQM(I),[JCO(J,I),J=1,IQM(I)] (BN,A5,12I5)`

  `BSET` - set equal to `LARGE` to indicate that large component basis is defined explicitly

  `IQM` - highest angular quantum number L plus one, e.g. s(1),p(2) ....

  `JCO` - number of blocks for each L-value

- *For each block read exponents and coefficients:*

  In the present version only uncontracted basis sets are used. The large component basis is given explicitly, whereas the small component basis is generated by the kinetic balance prescription. If contracted basis sets are to be used, the two components must each be specified explicitly. This will be possible in the next version of **DIRAC** .

  Card 3.2.1 `FRMT,NUC,NRC,ISGEN (BN,A,I4,2I5)`

  `FRMT` - format for reading of exponents and coefficients

  `H` - high precision. For each exponent the first line is read in format `(4F20.8)` and addtional lines containing only coefficients are read in format`(20X,3F20.8)`.

  `F` - free format

  `(blank)` - default precision. For each exponent the first line is read in format `(8F10.4)` and addtional lines containing only coefficients are read in format`(10X,7F10.4)`.

  `NUC` - number of primitve exponents

  `NRC` - number of contracted functions (set equal to zero for uncontracted basis)

  `ISGEN` - specification of how to generate small component functions by kinetic balance:

  - `ISGEN = 0`: No small component functions generated
  - `ISGEN = 1`: Small component functions generated upwards, e.g. $p \rightarrow d$
  - `ISGEN = 2`: Small component functions generated downwards, e.g. $p \rightarrow s$
  - `ISGEN = 3`: Small component functions generated both upwards and downwards, e.g. $p \rightarrow s, d$

  Card 3.2.2 `[ALPHA(K),[CPRIM(K,L),L=1,NRC],K=1,NUC]`

  Read exponents and coefficients. See card 3.2.1.

Card 4 `FINISH` - keyword to indicate end of file

Table A.1: Example definitions of the binary groups (- indicates blank character)

Group	NSYMOP	SYMOP	Operations
D_{2h}	3	`--Z--Y--X`	$\sigma_{xy}, \sigma_{xz}, \sigma_{yz}$
D_2	2	`XY--YZ---`	C_z^2, C_x^2
C_{2v}	2	`-Y-X-----`	σ_{xz}, σ_{yz}
C_{2h}	2	`--ZXYZ---`	σ_{xy}, i
C_2	1	`XY-------`	C_z^2
C_s	1	`--Z------`	σ_{xy}
C_i	1	`XYZ------`	i
C_1	0	`---------`	

Table A.2: Summary of basis file structure

Card	Input	Format
1:	`KEYWRD = 'INTGRL'`	`(A6)`
2a:	`TITLE(1)`	`(A72)`
2b:	`TITLE(2)`	`(A72)`
3:	`CRT,NONTYP,NSYMOP,[SYMOP(I),I=1,3],THRS`	`(BN,A1,I4,I5,3A3,1X,D10.2)`
	• For each atomic type:	`I = 1,NONTYP`
3.1:	`Q,NONT(I),QEXP`	`(BN,F10.0,I5,F20.5)`
	• For each symmetry independent center:	`J = 1,NONT(I)`
3.1.1:	`NAMN(NUCIND),[CORD(J,NUCIND), J = 1,3]`	`(BN,A4,3F20.0)`
3.2:	`BSET,IQM(I),[JCO(J,I),J=1,IQM(I)]`	`(BN,A5,12I5)`
	• For each block:	`J = 1,IQM(I)`
3.2.1:	`FRMT,NUC,NRC,ISGEN`	`(BN,A,I4,2I5)`
3.2.2:	`[ALPHA(K),[CPRIM(K,L),L=1,NRC],K=1,NUC]`	see text
4:	`KEYWRD = 'FINISH'`	`(A6)`

A.6 Output files

A formatted output file is conneted to **DIRAC** through standard output. The user may modify the output by setting various print levels in the menu file. **DIRAC** willl in addition produce a formatted file `DFCYCL` containing a summary of the SCF process. In addition to the formatted **DIRAC** produces a number of unformatted files. They are

- **Control files**

`DFDIIS`	information about DIIS process
`DFEVEC`	direct access file with DIIS error vector

- **Coefficients**

`DFCOEF`	MO-coefficients from current SCF-iteration
`DFCMOS`	coefficients from current SCF-iteration in MO-basis

- **One-electron integrals and matrices**

`DF1INT`	one-electron integrals contributing to one-electron Fock matrix
`DFOVLP`	overlap matrices
`DFTMAT`	MO-transformation matrix
`DFFCK1`	One-electron Fock matrix (in QO basis)

- **Two-electron integrals and matrices**
  For each integral class (`XX = LL,SL,SS`) in conventional mode:

`DFXXSA`	sorted singlet integrals (both Coulomb and exchange contributions)
`DFXXSB`	sorted singlet integrals (both Coulomb and exchange contributions)
`DFXXTA`	sorted triplet integrals (only exchange contributions)
`DFXXTB`	sorted triplet integrals (only exchange contributions)
`DFTWXX`	scratch file of unsorted integrals from **HERMIT**
`DFXXTB`	scratch file used in sorting process

In addition:

`DFFCK2`	two-electron Fock matrix in QO-basis

A.7 Restart

DIRAC has robust restart facilities:

- When calculating a **new point** on a potential surface, **DIRAC** can start from the coefficients (the file `DFCOEF`) , from the two-electron Fock matrix in AO-basis (the file `DFFCK2`) or from solutions of the one-electron Fock matrix (bare nucleus approximation). Default is to start from coefficients if the unformatted vector file `DFCOEF` is present; otherwise **DIRAC** uses the bare nucleus approxiamtion. Restart on Fock matrix may be specified by the keyword `.TRIFCK`.

- When restarting on the **same point** on the potential surface, **DIRAC** needs the formatted file typ DFCYCL to update status of the SCF process. The full SCF summary will be provided at the end of the current iterations, so that the output file from the previous SCF iterations is generally not needed. In addition **DIRAC** needs the coefficients (file `DFCOEF`). To restart on DIIS, **DIRAC** needs the following files: `DFDIIS, DFCMOS, DFFOCK` and `DFEVEC`. If DIIS is not requested, **DIRAC** may restart on damping if the file `DFFOCK` is present.

A.8 Test example

We consider the calculation of the F_2- molecule as a test case. Notice the use of the hash symbol `#}`to insert comments and turn off keywords.

A.8.1 Basis file

```
INTGRL
Fluorine molecule. Experimental bond length: 1.4178 A
Basis: F9s5p1d
C   1    3X  Y  Z  A   .10D-15
# D2h - symmetry is specified.
       9.0    1
F1      .000000000000000    .000000000000000     .70890000000000       *
LARGE    3    1    1    1
    9    0    3
 9994.7900
 1506.0300
 350.26900
# Comments may be inserted anywhere
 104.05300
 34.843200
 12.216400
 4.3688000
 1.2078000
  .3634000
    5    0    3
 44.355500
 10.082000
 2.9959000
  .9383000
  .2733000
    1    0    3
 1.6200000
FINISH
```

A.8.2 Menu file

```
*DIRAC
.TITLE
Fluorine molecule. Restr.kin.bal. Sph.tr.
.DIRECT
 1 1 1
.DHFCAL
.DHFANA
.NSYM
2
#.URKBAL
*READIN
*DHFCAL
.NELECT
 10 8
.PRINT
2
.DODAMP
0.25
.DODIIS
5000.0
#  The following three keywords specify that the large component
#  coefficients for all spinors are to be printed. In OFFSET and
#  SPINOR very large values are given. These will be modified in the
#  program down to maximum possible values, that is printing all
#  spinors.
.OFFSET
-100 -100
.SPINOR
 100 100
.VECPRI
 1 0
.EIGPRI
 1 1
*DHFANA
.MULPOP
*END OF
```

Appendix B

Reduction of the Breit term to non-relativistic form

We consider the Breit interaction

$$\hat{g}^{Breit} = -\left\{ \frac{\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{2r_{12}} + \frac{(\boldsymbol{\alpha}_1 \cdot \mathbf{r}_{12})\,(\boldsymbol{\alpha}_2 \cdot \mathbf{r}_{12})}{2r_{12}^3} \right\} \tag{B.1}$$

It can be reformulated to [168]

$$\hat{g}^{Breit} = \hat{g}^{Gaunt} + \hat{g}^{gauge} \tag{B.2}$$

The first term is then the Gaunt-term

$$\hat{g}^{Gaunt} = -\frac{\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} \tag{B.3}$$

and the second term is a gauge-dependent term

$$\hat{g}^{gauge} = -\frac{(\boldsymbol{\alpha}_1 \cdot \boldsymbol{\nabla}_1)\,(\boldsymbol{\alpha}_2 \cdot \boldsymbol{\nabla}_2)\, r_{12}}{2} \tag{B.4}$$

where $\boldsymbol{\nabla}_1$ and $\boldsymbol{\nabla}_2$ act *only* on r_{12} and not on the wave function. We consider the Foldy-Wouthuysen transformation of the Breit operator to order $(Z\alpha)^2$. In particular we are interested in the separate contributions of the Gaunt and the gauge-dependent term. This requires the evaluation of the anticommutator expression [40]

$$\frac{1}{4m^2c^2}\left[(\boldsymbol{\alpha}_2 \cdot \mathbf{p}_2)\,, \left[(\boldsymbol{\alpha}_1 \cdot \mathbf{p}_2)\,, \hat{g}^{Breit}\right]_+\right]_+ \tag{B.5}$$

which is quite a laborious task. Let us first outline the general strategy for the evaluation of the commutator expression. The commutator we want to evaluate may be written as

$$\left[\alpha_{2i} p_{2i}, \left[\alpha_{1j} p_{1j}, \alpha_{1k}\alpha_{2m} M_{km}\right]_+\right]_+ \tag{B.6}$$

where M_{km} is a function of $\mathbf{r}_{12}$ and symmetric with respect to particle exchange

$$M = M(\mathbf{r}_{12}) = M(-\mathbf{r}_{12}); \qquad \mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2 \tag{B.7}$$

The commutator consists of a space part, represented by momentum operators and coordinates, and a spin part, represented by Dirac $\boldsymbol{\alpha}$-matrices. The latter go into Pauli spin matrices in the two-component form. The spin part is best handled using anticommutator relations, in particular the relation

$$[\alpha_i, \alpha_j]_+ = 2\delta_{ij} \tag{B.8}$$

whereas the space part is best handled using commutator relations. We will therefore rearrange the commutator expression to a form that alows a more straightforward evaluation. We then employ the commutator relations

$$\begin{array}{lclcll} [AB, C] & = & A\,[B, C] & + & [A, C] & (a) \\ [A, BC] & = & [A, B]\,C & + & B\,[A, C] & (b) \\ [AB, C]_+ & = & [A, C]_+\,B & + & A\,[B, C] & (c) \\ [A, BC]_+ & = & [A, B]_+\,C & - & B\,[A, C] & (d) \end{array} \tag{B.9}$$

Consider first the inner anticommutator. Using Eq.B.9c it can be expanded as

$$[\alpha_{1j}p_{1j}, \alpha_{1k}\alpha_{2m}M_{km}]_+ = [\alpha_{1j}, \alpha_{1k}\alpha_{2m}M_{km}]_+\,p_{1j} + \alpha_{1j}\,[p_{1j}, \alpha_{1k}\alpha_{2m}M_{km}] \tag{B.10}$$

Using Eq.B.9d the first term is rearranged to

$$\begin{aligned} [\alpha_{1j}, \alpha_{1k}\alpha_{2m}M_{km}]_+\,p_{1j} &= [\alpha_{1j}, \alpha_{1k}]_+\,\alpha_{2m}M_{km}p_{1j} - \alpha_{1k}\,[\alpha_{1j}, \alpha_{2m}M_{km}]\,p_{1j} \\ &= 2\delta_{jk}\alpha_{2m}M_{km}p_{1j} \end{aligned} \tag{B.11}$$

and using Eq.B.9b the second term becomes

$$\begin{aligned} \alpha_{1j}\,[p_{1j}, \alpha_{1k}\alpha_{2m}M_{km}] &= \alpha_{1j}\,[p_{1j}, \alpha_{1k}\alpha_{2m}]\,M_{km} + \alpha_{1j}\alpha_{1k}\alpha_{2m}\,[p_{1j}, M_{km}] \\ &= \alpha_{1j}\alpha_{1k}\alpha_{2m}\,[p_{1j}, M_{km}] \end{aligned} \tag{B.12}$$

The two surviving terms are processed through the outer anticommutator using the same techniques and finally gives four terms

$$\begin{aligned} \left[\alpha_{2i}p_{2i}, [\alpha_{1j}p_{1j}, \alpha_{1k}\alpha_{2m}M_{km}]_+\right]_+ &= 4M_{ij}p_{1i}p_{2j} && (a) \\ &+ 2\alpha_{2i}\alpha_{2m}\,[p_{2i}, M_{jm}]\,p_{1j} && (b) \\ &+ 2\alpha_{1j}\alpha_{1k}\,[p_{1j}, M_{ki}]\,p_{2i} && (c) \\ &+ \alpha_{2i}\alpha_{2m}\alpha_{1j}\alpha_{1k}\,[p_{2i}, [p_{1j}, M_{km}]] && (d) \end{aligned} \tag{B.13}$$

The final term can be simplified by noting that

$$[p_{2i},[p_{1j},M_{km}]]=(p_{2i}p_{1j}M_{km}) \tag{B.14}$$

where the parenthesis indicate that $p_{2i}p_{1j}$ act on M_{km} and not on the wave function. Products of $\boldsymbol{\alpha}$ - terms are handled by the relation

$$\alpha_i\alpha_j=\delta_{ij}+i\epsilon_{ijk}\alpha_k \tag{B.15}$$

Consider the **Gaunt term**. We make the identification

$$M_{km}=-\delta_{km}r_{12}^{-1} \tag{B.16}$$

which means that we need the relations

$$\begin{aligned}\left[p_{1i},r_{12}^{-1}\right] &= ir_{12}^{-3}r_{12i}\\ \left[p_{2i},r_{12}^{-1}\right] &= -ir_{12}^{-3}r_{12i}\\ p_{2i}p_{1j}r_{12}^{-1} &= 3r_{12i}r_{12j}r_{12}^{-5}-(4\pi/3)\delta(\mathbf{r}_{12})\delta_{ij}-\delta_{ij}r_{12}^{-3}\end{aligned} \tag{B.17}$$

The latter relation may be compared to Eq.(2.96) in Moss [40]. The contribution from the Gaunt-term to the Breit-Pauli Hamiltonian is therefore

$$\begin{aligned}\hat{g}^{Gaunt}: \quad &- \frac{1}{m^2c^2r_{12}}\left(\mathbf{p}_1\cdot\mathbf{p}_2\right) && (a)\\ &- \frac{1}{2m^2c^2r_{12}^3}\left[\boldsymbol{\alpha}_2\cdot\left(\mathbf{r}_{12}\times\mathbf{p}_1\right)+\left(\mathbf{r}_{12}\cdot\nabla_1\right)\right] && (b)\\ &+ \frac{1}{2m^2c^2r_{12}^3}\left[\boldsymbol{\alpha}_1\cdot\left(\mathbf{r}_{12}\times\mathbf{p}_2\right)+\left(\mathbf{r}_{12}\cdot\nabla_2\right)\right] && (c)\\ &- \frac{1}{4m^2c^2}\left(\boldsymbol{\alpha}_2\cdot\mathbf{p}_2\right)\left(\boldsymbol{\alpha}_1\cdot\mathbf{p}_1\right)\left(\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2\right)r_{12}^{-1} && (d)\end{aligned} \tag{B.18}$$

Further processing and replacing $\boldsymbol{\alpha}$ with $\boldsymbol{\sigma}$ gives

$$\begin{aligned}\hat{g}^{Gaunt}: \quad &- \frac{1}{m^2c^2r_{12}}\left(\mathbf{p}_1\cdot\mathbf{p}_2\right) && (a)\\ &+ \frac{1}{2m^2c^2r_{12}^3}\left[\boldsymbol{\sigma}_1\cdot\left(\mathbf{r}_{12}\times\mathbf{p}_2\right)-\boldsymbol{\sigma}_2\cdot\left(\mathbf{r}_{12}\times\mathbf{p}_1\right)\right] && (b)\\ &+ \frac{1}{4m^2c^2}\left[r_{12}^{-3}\left(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\right)-3r_{12}^{-5}\left(\boldsymbol{\sigma}_1\cdot\mathbf{r}_{12}\right)\left(\boldsymbol{\sigma}_2\cdot\mathbf{r}_{12}\right)-(8\pi/3)\left(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\right)\delta(\mathbf{r}_{12})\right] && (c)\\ &- \frac{1}{2m^2c^2r_{12}^3}\left(\mathbf{r}_{12}\cdot\boldsymbol{\nabla}_{12}\right) && (d)\\ &- \frac{1}{m^2c^2}\pi\delta(\mathbf{r}_{12}) && (e)\end{aligned}$$

(B.19)

For **gauge-dependent** term we find

$$M_{km} = -\frac{1}{2}\nabla_{1k}\nabla_{2m}r_{12} \tag{B.20}$$

where we again note $\boldsymbol{\nabla}_1$ and $\boldsymbol{\nabla}_2$ act only on r_{12} and not on the wave function. The contribution from the gauge-dependent term to the Breit-Pauli Hamiltonian is therefore

$$\begin{aligned}
\hat{g}^{gauge}: \quad & - \frac{1}{2m^2c^2}\nabla_{2i}\nabla_{1j}r_{12}p_{1j}p_{2i} && (a)\\
& + \frac{1}{2m^2c^2r_{12}^3}\left(\mathbf{r}_{12}\cdot\boldsymbol{\nabla}_1\right) && (b)\\
& - \frac{1}{2m^2c^2r_{12}^3}\left(\mathbf{r}_{12}\cdot\boldsymbol{\nabla}_2\right) && (c)\\
& + \frac{1}{8m^2c^2}\left(\boldsymbol{\alpha}_2\cdot\boldsymbol{\nabla}_2\right)\left(\boldsymbol{\alpha}_2\cdot\boldsymbol{\nabla}_2\right)\left(\boldsymbol{\alpha}_1\cdot\boldsymbol{\nabla}_1\right)\left(\boldsymbol{\alpha}_1\cdot\boldsymbol{\nabla}_1\right)r_{12} && (d)
\end{aligned} \tag{B.21}$$

The final term may be collapsed into

$$\frac{1}{8m^2c^2}\nabla_1^2\nabla_2^2 r_{12} \tag{B.22}$$

Further processing and replacing $\boldsymbol{\alpha}$ with $\boldsymbol{\sigma}$ gives

$$\begin{aligned}
\hat{g}^{gauge}: \quad & - \frac{1}{2m^2c^2}\left(\mathbf{p}_1\cdot\boldsymbol{\nabla}_1\right)\left(\mathbf{p}_2\cdot\boldsymbol{\nabla}_2\right)r_{12} && (a)\\
& + \frac{1}{2m^2c^2r_{12}^3}\left(\mathbf{r}_{12}\cdot\boldsymbol{\nabla}_{12}\right) && (b)\\
& + \frac{1}{m^2c^2}\pi\delta(\mathbf{r}_{12}) && (c)
\end{aligned} \tag{B.23}$$

The reduced Breit term is obtained by combining Eq.(B.19) and Eq.(B.23):

$$\begin{aligned}
\hat{g}^{Breit}: \quad & \frac{1}{m^2c^2}\left[r_{12}^{-1}\left(\mathbf{p}_1\cdot\mathbf{p}_2\right)+\frac{1}{2}\left(\mathbf{p}_1\cdot\boldsymbol{\nabla}_1\right)\left(\mathbf{p}_2\cdot\boldsymbol{\nabla}_2\right)r_{12}\right] && (a)\\
+ \quad & \frac{1}{2m^2c^2r_{12}^3}\left[\boldsymbol{\sigma}_1\cdot\left(\mathbf{r}_{12}\times\mathbf{p}_2\right)-\boldsymbol{\sigma}_2\cdot\left(\mathbf{r}_{12}\times\mathbf{p}_1\right)\right] && (b)\\
+ \quad & \frac{1}{4m^2c^2}\left[r_{12}^{-3}\left(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\right)-3r_{12}^{-5}\left(\boldsymbol{\sigma}_1\cdot\mathbf{r}_{12}\right)\left(\boldsymbol{\sigma}_2\cdot\mathbf{r}_{12}\right)-(8\pi/3)\left(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\right)\delta(\mathbf{r}_{12})\right] && (c)
\end{aligned} \tag{B.24}$$

The first term in Eq.(B.24) is the orbit-orbit interaction in the form given by Huang [146], where $\boldsymbol{\nabla}_1$ and $\boldsymbol{\nabla}_2$ operate only on r_{12} and not on the wave function. It is straightforwardly rearranged to its more familiar form

$$\frac{1}{2m^2c^2}\left[r_{12}^{-1}\left(\mathbf{p}_1\cdot\mathbf{p}_2\right)+r_{12}^{-3}\left(\mathbf{p}_1\cdot\mathbf{r}_{12}\right)\left(\mathbf{p}_2\cdot\mathbf{r}_{12}\right)\right] \tag{B.25}$$

Note that the two forms of the orbit-orbit interaction corresponds to the two forms of the Breit term.

Appendix C

Symmetry in relativistic systems

C.1 General overview

THESES NOTES ARE BASED ON TWO LECTURES HELD AT THE UNIVERSITY OF TROMSØ DECEMBER 1994.

The full symmetry group of a dynamical system consists of all operators that commute with its Hamiltonian. Symmetry operations are unitary or antiunitary. Here we will only consider unitary symmetry operations acting on spatial and spin coordinates of the system. The non-relativistic Hamiltonian has the general form

$$\hat{H} = \hat{T} + \hat{V} \tag{C.1}$$

The kinetic energy operator $\hat{T}$ is totally symmetric under any symmetry operation, whereas the potential energy operator $\hat{V}$ determines the symmetry of the system. In molecular systems the translation of the center of mass is separated out, so that our discussion of symmetry will be limited to point group operations that keep at least one point fixed. We may write a general point group symmetry operator as

$$\hat{G} = \hat{G}^{\mathbf{r}}\left(\phi_{\mathbf{r}}, \mathbf{n}_{\mathbf{r}}, p_{\mathbf{r}}\right) \hat{G}^{\boldsymbol{\eta}}\left(\phi_{\eta}, \mathbf{n}_{\eta}, p_{\eta}\right); \qquad p_i = 0, 1 \tag{C.2}$$

where $\hat{G}^{\mathbf{r}}\left(\phi_{\mathbf{r}}, \mathbf{n}_{\mathbf{r}}, p_{\mathbf{r}}\right)$ and $\hat{G}^{\boldsymbol{\eta}}\left(\phi_{\eta}, \mathbf{n}_{\eta}, p_{\eta}\right)$ act on spatial $\mathbf{r}$ and spin η coordinates, respectively. They have the form

$$\hat{G}\left(\phi, \mathbf{n}, p\right) = \hat{i}^{p} \hat{R}\left(\phi, \mathbf{n}\right) \quad , \quad (p = 0, 1) \tag{C.3}$$

where $\hat{i}$ represent inversion and $\hat{R}\left(\phi, \mathbf{n}\right)$ is a rotation ϕ about an axis given by the unit vector **n**. Special cases are

$$\begin{array}{lll} \text{Rotation:} & \hat{G}\left(\phi,\mathbf{n},0\right) & \\ \text{Inversion:} & \hat{G}\left(0,\mathbf{n},1\right) & \text{(arbitrary } \mathbf{n}\text{)} \\ \text{Reflection:} & \hat{G}\left(\pi,\mathbf{n},1\right) & \\ \text{Improper rotation:} & \hat{G}\left(\phi,\mathbf{n},1\right) & \end{array} \tag{C.4}$$

The action of a symmetry operation on a function is determined by its action on the coordinates

$$\hat{G}\psi(\hat{G}(\mathbf{r},\eta)) = \psi(\mathbf{r},\eta) \quad \Rightarrow \quad \hat{G}\psi(\mathbf{r},\eta) = \psi(\hat{G}^{-1}(\mathbf{r},\eta)) \tag{C.5}$$

The non-relativistic Hamiltonian is spin-free, which allows a separate handling of spin and spatial symmetry. The spin may be integrated out of equations and one may form spin-free functions adapted to point group symmetry. In the relativistic domain the spin and spatial degrees of freedom are coupled and no such separation is possible. In a molecular system with an even(odd) number of electrons, the eigenfunctions of the electronic Hamiltonian have integral(half-integral) spin and may be denoted *boson* (*fermion*) functions. A striking difference between boson and fermion functions, is that the latter type functions change sign under a rotation 2π about an arbitrary axis, which has in fact been verified experimentally in both neutron and NMR interfermometry experiments. For boson functions a rotation 2π correspond to the identity operation.

The symmetry of fermion functions is usually handled using *double groups*. Double groups are introduced by adding an extra element $\overline{E}$ representing a rotation 2π about an arbitrary axis and therefore commuting with all symmetry operations. By this 'trick' it is possible to recover all the results of standard group theory. In these notes we shall derive explicit representations of the various point group operations in spatial and spin coordinates. We shall see that the resulting representations for rotations are completely congruent with what is obtained from double group theory. For improper rotations involving, there is a decisive difference that requires further exploration.

C.2 Notation

In what follows we employ the following notation

- the *Einsteins summation convention*: a repeated index is taken to mean summation over all possible values of the index
- the *three-dimensional Levi-Cevita symbol*

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if ijk is an even permutation of 1,2,3} \\ -1 & \text{if ijk is an odd permutation of 1,2,3} \\ 0 & \text{for all other cases} \end{cases} \tag{C.6}$$

We note the relation $\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$

Using the above notation scalar and vectors products may be expressed as

$$\mathbf{A} \cdot \mathbf{B} = A_i B_i; \qquad \mathbf{A} \times \mathbf{B} = \mathbf{e}_i \epsilon_{ijk} A_j B_k \tag{C.7}$$

where $\mathbf{e}_i$ are unit vectors along the main coordinate axes. Alternativelym, a vector product can be written as a 3 × 3 determinant

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \tag{C.8}$$

from which we see that a general 3 × 3 determinant can be expressed in terms of the Levi-Cevita symbol

$$\det(\mathbf{A}) = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} = \epsilon_{ijk} A_{1i} A_{2j} A_{3k} \tag{C.9}$$

We may generalize to

$$\epsilon_{ijk} A_{li} A_{mj} A_{nk} = \epsilon_{lmn} \det(\mathbf{A}) \tag{C.10}$$

C.3 Rotations

C.3.1 Rotation about main axes

A rotation ω of a function about the z axis correspond to a rotation $-\omega$ of coordinates:

$$\begin{aligned} x' &= r\sin\theta\cos(\phi-\omega) &&= r\sin\theta\cos\phi\cos\omega + r\sin\theta\sin\phi\sin\omega \\ &= x\cos\omega + y\sin\omega && \\ y' &= r\sin\theta\sin(\phi-\omega) &&= -r\sin\theta\cos\phi\sin\omega + r\sin\theta\sin\phi\cos\omega \\ &= -x\sin\omega + y\cos\omega && \\ z' &= r\cos\theta && \\ &= z && \end{aligned} \tag{C.11}$$

Consider an infinitesimal rotation $d\omega$ about the z-axis. We use the relations

$$\sin d\omega \approx d\omega \quad , \quad \cos d\omega \approx 1 \tag{C.12}$$

which gives

$$d\mathbf{r}' = \left(x' - x, y' - y, z' - z\right) = (y, -x, 0)\, d\omega \tag{C.13}$$

The resulting function may be written as a first order Taylor expansion :

$$\psi(\mathbf{r}^{'}) \approx \psi(\mathbf{r}) + d\mathbf{r}\cdot\nabla\psi(\mathbf{r}) \ \ = [1 + d\omega(y\frac{d}{dx} - x\frac{d}{dy})]\psi(\mathbf{r}) \ \ = [1 - id\omega\hat{l}_z]\psi(\mathbf{r}) \qquad \text{(C.14)}$$

From the equivalence of the main axes we conclude

$1 - d\omega\hat{l}_x$	infinitesimal rotation about the x-axis
$1 - d\omega\hat{l}_y$	infinitesimal rotation about the y-axis
$1 - d\omega\hat{l}_z$	infinitesimal rotation about the z-axis

(C.15)

C.3.2 Rotation about arbitrary axis

We consider an infinitesimal rotation $d\phi$ about an arbitrary axis identified by the unit vector $\mathbf{n}$. We employ the notation $\hat{R}\,(d\phi,\mathbf{n})$ to describe this rotation. By geometric construction we find

$$d\mathbf{r} = -d\phi\,(\mathbf{n}\times\mathbf{r}) \qquad \text{(C.16)}$$

Insertion in a first order Taylor expansion of the rotated function gives

$$\psi(\mathbf{r}^{'}) \approx \psi(\mathbf{r}) + d\mathbf{r}\cdot\nabla\psi(\mathbf{r}) = \psi(\mathbf{r}) - d\phi\,(\mathbf{n}\times\mathbf{r})\cdot\nabla\psi(\mathbf{r}) = [1 - d\phi\mathbf{n}\cdot(\mathbf{r}\times\nabla)]\,\psi(\mathbf{r}) \qquad \text{(C.17)}$$

This allows the identification

$$\hat{R}\,(d\phi,\mathbf{n}) = 1 - id\phi\left(\mathbf{n}\cdot\hat{\mathbf{l}}\right) \qquad \text{(C.18)}$$

where $\hat{\mathbf{l}}$ is the angular momentum operator (in atomic units)

$$\hat{\mathbf{l}} = \hat{\mathbf{r}}\times\hat{\mathbf{p}} = -i\,(\mathbf{r}\times\nabla) \qquad \text{(C.19)}$$

The operator for a finite rotation is generated from infinitesimal rotations

$$\hat{R}\,(\phi,\mathbf{n}) = \lim_{k\to\infty}\left[\hat{R}\left(\frac{\phi}{k},\mathbf{n}\right)\right]^k = \lim_{k\to\infty}\left[1 - i\frac{\phi}{k}\left(\mathbf{n}\cdot\hat{\mathbf{l}}\right)\right]^k = e^{-i(\mathbf{n}\cdot\hat{\mathbf{l}})} \qquad \text{(C.20)}$$

C.3.3 Angular momentum — a short repetition

In the previous section we have seen that the operators for angular momentum are generators for infinitesimal rotations about the main axes. We therefore give a short summary of angular momentum theory.

Genral angular momentum operators are defined from the commutation relations

$$\left[\hat{j}_x, \hat{j}_y\right] = i\hat{j}_z \quad , \quad \left[\hat{j}_y, \hat{j}_z\right] = i\hat{j}_x \quad , \quad \left[\hat{j}_z, \hat{j}_x\right] = i\hat{j}_y \tag{C.21}$$

The commutation relation may be written in a more compact form by

$$\hat{\mathbf{j}} \times \hat{\mathbf{j}} = i\hat{\mathbf{j}} \tag{C.22}$$

We denote eigenfunctions of angular momentum by ψ_{jm}. The following relations apply:

$$\begin{aligned} \hat{j}^2\psi_{jm} &= j(j+1)\psi_{jm} \\ \hat{j}_z\psi_{jm} &= m\psi_{jm} \end{aligned} \tag{C.23}$$

We introduce ladder operators $\hat{j}_+$ and $\hat{j}_-$

$$\hat{j}_\pm = \hat{j}_x \pm i\hat{j}_y \tag{C.24}$$

whose operation on

$$\hat{j}_\pm\psi_{jm} = \sqrt{(j(j+1) - m(m\pm 1))}\psi_{j,m\pm 1} \tag{C.25}$$

For later use we note the relations:

$$j_x = \frac{1}{2}\left(j_+ + J_-\right) \quad , \quad j_y = -i\frac{1}{2}\left(j_+ - j_-\right) \tag{C.26}$$

C.3.4 Matrix representations of rotation operators

The product of two rotations is a rotation. The set of rotation operators

$$\hat{R}\left(\phi, \mathbf{n}\right) = e^{-i\phi(\mathbf{n}\cdot\hat{\mathbf{j}})} \tag{C.27}$$

form a continuous group, the full rotation group R_3. Irreducible representations of the full rotation group are labeled by j and are $(2j+1)$-fold degenerate. Matrix representations of $\hat{\mathbf{j}}$ and $\hat{R}\left(\phi, \mathbf{n}\right)$ may be constructed in a basis $\{\psi_{jm}\}$ for any j. In this section we consider matrix representations of $j = \frac{1}{2}$ og $j = 1$.

Case 1: $j = \frac{1}{2}$

Basis for $j = \frac{1}{2}$ are the spin functions α and β and we set $\mathbf{j} = \mathbf{s}$. We have the relations

$$\begin{aligned} &s_z\alpha = \tfrac{1}{2}\alpha \quad && s_z\beta = -\tfrac{1}{2}\beta \\ &s_+\alpha = 0 && s_+\beta = \alpha \\ &s_-\alpha = \beta && s_-\beta = 0 \end{aligned} \tag{C.28}$$

from which we obtain the matrix representations

$$s_z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad , \quad s_+ = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad , \quad s_- = \frac{1}{2}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{C.29}$$

Using (C.26) we obtain

$$s_x = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad s_y = \frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{C.30}$$

In sum we have

$$\mathbf{s} = \frac{1}{2}\boldsymbol{\sigma} \tag{C.31}$$

where $\boldsymbol{\sigma}$ are *Pauli spin matrices.* We note the property

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k = \begin{cases} 1 & , \quad i = j \\ i\epsilon_{ijk}\sigma_k & , i \neq j \end{cases} \tag{C.32}$$

Important relations are

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B}) \tag{C.33}$$

and

$$\begin{aligned} &(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B})(\boldsymbol{\sigma}\cdot\mathbf{C}) \\ &\quad = i\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) + (\boldsymbol{\sigma}\cdot\mathbf{A})(\mathbf{B}\cdot\mathbf{C}) - (\boldsymbol{\sigma}\cdot\mathbf{B})(\mathbf{A}\cdot\mathbf{C}) + (\boldsymbol{\sigma}\cdot\mathbf{C})(\mathbf{A}\cdot\mathbf{B}) \end{aligned} \tag{C.34}$$

Matrix representaitons of the rotation operators for $j = \frac{1}{2}$ is obtained from

$$\hat{R}^{\frac{1}{2}}(\phi, \mathbf{n}) = e^{-i\phi(\mathbf{n}\cdot\hat{\mathbf{s}})} \to \mathbf{R}^{\frac{1}{2}}(\phi, \mathbf{n}) = e^{-i\frac{1}{2}\phi(\mathbf{n}\cdot\boldsymbol{\sigma})} = \sum_{m=0}(-i)^m \frac{\left(\frac{1}{2}\phi\right)^m}{m!}(\mathbf{n}\cdot\boldsymbol{\sigma})^m \tag{C.35}$$

Considerable simplification is obtained by noting that

$$(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = n_j\sigma_j n_k\sigma_k = n_i n_i + i\epsilon_{ijk}\sigma_i n_j n_k = 1 + i\boldsymbol{\sigma}(\mathbf{n}\times\mathbf{n}) = 1 \tag{C.36}$$

which implies

$$(\mathbf{n}\cdot\boldsymbol{\sigma})^{2m} = 1 \quad , \quad (\mathbf{n}\cdot\boldsymbol{\sigma})^{2m+1} = (\mathbf{n}\cdot\boldsymbol{\sigma}) \tag{C.37}$$

We may therefore write

$$\begin{aligned} \mathbf{R}^{\frac{1}{2}}(\phi, \mathbf{n}) &= \textstyle\sum_{n=0}(-i)^n \frac{\left(\frac{1}{2}\phi\right)^n}{n!}(\mathbf{n}\cdot\boldsymbol{\sigma})^n \\ &= \textstyle\sum_{n=0}(-i)^{2n} \frac{\left(\frac{1}{2}\phi\right)^{2n}}{(2n)!}(\mathbf{n}\cdot\boldsymbol{\sigma})^{2n} + \sum_{n=0}(-i)^{(2n+1)} \frac{\left(\frac{1}{2}\phi\right)^{(2n+1)}}{(2n+1)!}(\mathbf{n}\cdot\boldsymbol{\sigma})^{(2n+1)} \\ &= \textstyle\sum_{n=0}(-1)^{2n} \frac{\left(\frac{1}{2}\phi\right)^{2n}}{(2n)!} - i(\mathbf{n}\cdot\boldsymbol{\sigma})\sum_{n=0}(-1)^{2n} \frac{\left(\frac{1}{2}\phi\right)^{(2n+1)}}{(2n+1)!} \\ &= \cos\tfrac{1}{2}\phi - i(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\tfrac{1}{2}\phi \end{aligned} \tag{C.38}$$

Case 2: $j = 1$

As basis for $j = 1$ we may use the spherical harmonic functions

$$\begin{array}{lclcl} Y_{1,1} & = & -\left(\frac{3}{8}\pi\right)\sin\theta e^{i\phi} & \propto & -\frac{1}{\sqrt{2}}\left(x+iy\right) \\ Y_{1,0} & = & \left(\frac{3}{4}\pi\right)\cos\phi & \propto & z \\ Y_{1,-1} & = & \left(\frac{3}{8}\pi\right)\sin\theta e^{-i\phi} & \propto & \frac{1}{\sqrt{2}}\left(x-iy\right) \end{array} \tag{C.39}$$

In the basis of these functions the angular momentum operator is represented by

$$l_x = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad , \quad l_y = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad , \quad l_z = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{C.40}$$

We shall, however, use a basis of Cartesian coordinates ($q_1 = x$,$q_2 = y$ og $q_3 = z$) to generate representations of the angular momentum and rotation operators. som basis. The angular momentum operators may be written as

$$\hat{\mathbf{l}} = -i\left(\mathbf{r}\times\nabla\right) = -i\mathbf{e}_i\epsilon_{ijk}q_j\frac{\delta}{\delta q_k} \tag{C.41}$$

so that elements of representation matrices in the Cartesian basis is given by

$$\begin{array}{rclcl} \left\langle q_j \left|\hat{\mathbf{l}}\right| q_k\right\rangle & = & -i\mathbf{e}_i\epsilon_{ilm}\left\langle q_j \left|q_l\frac{\delta}{\delta q_m}\right| q_k\right\rangle & = & -i\mathbf{e}_i\epsilon_{ilm}\delta_{km}\left\langle q_j \mid q_l\right\rangle \\ & = & -i\mathbf{e}_i\epsilon_{ilm}\delta_{km}\delta_{jl}w & = & -i\mathbf{e}_i\epsilon_{ijk} \end{array} \tag{C.42}$$

The full representation matrices can now be written as

$$\tau_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad , \quad \tau_y = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \quad , \quad \tau_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{C.43}$$

In the coordinate basis the rotation operators are given by

$$\hat{R}\left(\phi,\mathbf{n}\right) = e^{-i\phi(\mathbf{n}\cdot\hat{\mathbf{j}})} \rightarrow \mathbf{R}^1\left(\phi,\mathbf{n}\right) = e^{-i\phi(\mathbf{n}\cdot\boldsymbol{\tau})} = \sum_{m=0}(-i)^m\frac{(\phi)^m}{m!}\left(\mathbf{n}\cdot\boldsymbol{\tau}\right)^m \tag{C.44}$$

In order to obtain a matrix representation of $\hat{R}^1$ we use the relations

$$\begin{array}{lclcl} (\mathbf{n}\cdot\boldsymbol{\tau})_{jk} & = & -in_i\epsilon_{ijk} & & \\ (\mathbf{n}\cdot\boldsymbol{\tau})^2_{jm} & = & -n_in_l\epsilon_{ijk}\epsilon_{lkm} & = & -n_in_l\left(\delta_{im}\delta_{jl}-\delta_{il}\delta_{jm}\right) \\ & = & -n_jn_m+n_in_i\delta_{jm} & = & \delta_{jm}-n_jn_m \\ (\mathbf{n}\cdot\boldsymbol{\tau})^3_{lm} & = & -in_k\epsilon_{klj}\left(\delta_{jm}-n_jn_m\right) & = & -in_k\epsilon_{klm}+in_kn_jn_m\epsilon_{klj} \\ & = & -in_k\epsilon_{klm} & = & (\mathbf{n}\cdot\boldsymbol{\tau})_{lm} \end{array} \tag{C.45}$$

We may conclude

$$(\mathbf{n}\cdot\boldsymbol{\tau})^{2m} = (\mathbf{n}\cdot\boldsymbol{\tau})^2 = I - \mathbf{n}\otimes\mathbf{n} \quad , \quad (\mathbf{n}\cdot\boldsymbol{\tau})^{(2m+1)} = (\mathbf{n}\cdot\boldsymbol{\tau}) \tag{C.46}$$

Note that the relation $(\mathbf{n}\cdot\boldsymbol{\tau})^3 = (\mathbf{n}\cdot\boldsymbol{\tau})$ does not imply $(\mathbf{n}\cdot\boldsymbol{\tau})^2 = 1$ sonce the matrix $(\mathbf{n}\cdot\boldsymbol{\tau})$ is singular. We now write

$$\begin{aligned}
\mathbf{R}^1(\phi,\mathbf{n}) &= \textstyle\sum_{n=0}(-i)^n\frac{(\phi)^n}{n!}(\mathbf{n}\cdot\boldsymbol{\tau})^n \\
&= 1+\textstyle\sum_{n=1}(-i)^{2n}\frac{(\phi)^{2n}}{(2n)!}(\mathbf{n}\cdot\boldsymbol{\tau})^{2n}+\sum_{n=0}(-i)^{(2n+1)}\frac{(\phi)^{(2n+1)}}{(2n+1)!}(\mathbf{n}\cdot\boldsymbol{\tau})^{(2n+1)} \\
&= 1+(\mathbf{n}\cdot\boldsymbol{\tau})^2\textstyle\sum_{n=1}(-1)^2\frac{(\phi)^{2n}}{(2n)!}+i(\mathbf{n}\cdot\boldsymbol{\tau})\sum_{n=0}(-1)^{2n}\frac{\left(\frac{1}{2}\phi\right)^{(2n+1)}}{(2n+1)!} \\
&= 1+(1-\cos\phi)(\mathbf{n}\cdot\boldsymbol{\tau})^2+i(\mathbf{n}\cdot\boldsymbol{\tau})\sin\phi
\end{aligned} \tag{C.47}$$

Using the trigonometric identities

$$\begin{aligned}
\cos^2\tfrac{1}{2}\phi+\sin^2\tfrac{1}{2}\phi &= 1 \\
\cos^2\tfrac{1}{2}\phi-\sin^2\tfrac{1}{2}\phi &= \cos\phi
\end{aligned} \tag{C.48}$$

this is simplified to

$$\mathbf{R}^1(\phi,\mathbf{n}) = 1+i(\mathbf{n}\cdot\boldsymbol{\tau})\sin\phi-2(\mathbf{n}\cdot\boldsymbol{\tau})^2\sin^2\frac{1}{2}\phi \tag{C.49}$$

C.3.5 Homomorphism between SO(3) and SU(2)

Consider the transformation of the Pauli spin matrices under the rotation operators

$$\hat{R}(\phi,\mathbf{n})\,\sigma_j\hat{R}^\dagger(\phi,\mathbf{n}) = \hat{R}(\phi,\mathbf{n})\,\sigma_j\hat{R}(-\phi,\mathbf{n}) = \sigma_i A_{ij}(\phi,\mathbf{n}) \tag{C.50}$$

This is best done by considering the transformation

$$\begin{aligned}
&\hat{R}(\phi,\mathbf{n})(\boldsymbol{\sigma}\cdot\mathbf{m})\,\hat{R}(-\phi,\mathbf{n}) \\
&\quad= \left[\cos\tfrac{1}{2}\phi-i(\boldsymbol{\sigma}\cdot\mathbf{n})\sin\tfrac{1}{2}\phi\right](\boldsymbol{\sigma}\cdot\mathbf{m})\left[\cos\tfrac{1}{2}\phi+i(\boldsymbol{\sigma}\cdot\mathbf{n})\sin\tfrac{1}{2}\phi\right] \\
&\quad= (\boldsymbol{\sigma}\cdot\mathbf{m})\cos^2\tfrac{1}{2}\phi+i\left[(\boldsymbol{\sigma}\cdot\mathbf{m}),(\boldsymbol{\sigma}\cdot\mathbf{n})\right]\cos\tfrac{1}{2}\phi\sin\tfrac{1}{2}\phi+(\boldsymbol{\sigma}\cdot\mathbf{n})(\boldsymbol{\sigma}\cdot\mathbf{m})(\boldsymbol{\sigma}\cdot\mathbf{n})\sin^2\tfrac{1}{2}\phi \\
&\quad= (\boldsymbol{\sigma}\cdot\mathbf{m})\cos^2\tfrac{1}{2}\phi+\boldsymbol{\sigma}\cdot(\mathbf{n}\times\mathbf{m})\sin\phi+\left[2(\boldsymbol{\sigma}\cdot\mathbf{n})(\mathbf{m}\cdot\mathbf{n})-(\boldsymbol{\sigma}\cdot\mathbf{m})\right]\sin^2\tfrac{1}{2}\phi \\
&\quad= (\boldsymbol{\sigma}\cdot\mathbf{m})+\boldsymbol{\sigma}\cdot(\mathbf{n}\times\mathbf{m})\sin\phi+\left[(\boldsymbol{\sigma}\cdot\mathbf{n})(\mathbf{m}\cdot\mathbf{n})-(\boldsymbol{\sigma}\cdot\mathbf{m})\right]2\sin^2\tfrac{1}{2}\phi
\end{aligned} \tag{C.51}$$

Setting $\mathbf{m} = \mathbf{e}_k \delta_{jk}$ the transformation of individual σ_j can be singled out. After reindexation we find

$$A_{jk}(\phi, \mathbf{n}) = \sigma_j \delta_{jk} + \sigma_j n_i \epsilon_{ijk} + \sigma_j (n_j n_k - \delta_{jk}) \, 2 \sin^2 \frac{1}{2}\phi \tag{C.52}$$

Comparison with (C.49) shows

$$A_{jk}(\phi, \mathbf{n}) = R_{jk}(\phi, \mathbf{n}) \tag{C.53}$$

It can be shown that the unitary transformation matrices $\mathbf{R}^{\frac{1}{2}}(\phi, \mathbf{n})$ have determinant 1 and form the special unitary group SU(2), whereas the matrices $\mathbf{R}^1(\phi, \mathbf{n})$ form the group SO(3) of orthogonal matrices with determinant 1. The transformation of Pauli spin matrices under elements of SU(2) provide a mapping of SU(2) into SO(3). The mapping is straighforwardly shown to be a homomorphism, but it is not single-valued since we have

$$\mathbf{R}^{\frac{1}{2}}((\phi + 2\pi), \mathbf{n}) = -\mathbf{R}^{\frac{1}{2}}(\phi, \mathbf{n}) \tag{C.54}$$

whereas

$$\mathbf{R}^1((\phi + 2\pi), \mathbf{n}) = \mathbf{R}^1(\phi, \mathbf{n}) \tag{C.55}$$

We see that rotations represented by $\mathbf{R}^1(\phi, \mathbf{n})$ have periodicity 2π, whereas rotations represented by $\mathbf{R}^{\frac{1}{2}}(\phi, \mathbf{n})$ have periodicity 4π. This is of consequence when we consider irreducible representations.

C.3.6 Direct product basis

The spin functions α and β form a basis for rotations $j = \frac{1}{2}$. A basis for $j = 1$ (and $j = 0$) in terms of spin functions is generated by forming the direct product

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \otimes \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{bmatrix} \tag{C.56}$$

The resulting functions are however, not all eigenfunctions of $\hat{j}$ og $\hat{j}_z$, but this can be corrected by the transformation

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{bmatrix} = \begin{bmatrix} \alpha\alpha \\ \frac{1}{\sqrt{2}}(\alpha\beta + \beta\alpha) \\ \beta\beta \\ \frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha) \end{bmatrix} = \begin{bmatrix} |1,1\rangle \\ |1,0\rangle \\ |1,-1\rangle \\ |0,0\rangle \end{bmatrix} \tag{C.57}$$

This gives the three components of a triplet in addition to a singlet function, which is what is expected from the coupling of two spins. We may construct a spin analogue to the Cartesian basis(C.39) as well by the transformation

$$\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 0 & 0 & 1 \\ i & 0 & 0 & i \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}\begin{bmatrix} \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}}(\alpha\alpha-\beta\beta) \\ \frac{i}{\sqrt{2}}(\alpha\alpha+\beta\beta) \\ \frac{1}{\sqrt{2}}(\alpha\beta+\beta\alpha) \\ \frac{1}{\sqrt{2}}(\alpha\beta-\beta\alpha) \end{bmatrix} \tag{C.58}$$

We may form corresponding direct products of the symmetry elements, e.g. a rotation ϕ about the z -axis:

$$\begin{bmatrix} e^{-i\frac{1}{2}\phi} & 0 \\ 0 & e^{i\frac{1}{2}\phi} \end{bmatrix} \otimes \begin{bmatrix} e^{-i\frac{1}{2}\phi} & 0 \\ 0 & e^{i\frac{1}{2}\phi} \end{bmatrix} = \begin{bmatrix} e^{-i\phi} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\phi} \end{bmatrix} \tag{C.59}$$

By transforming to the coordinate (Cartesian) basis (C.58) we obtain as expected

$$\begin{bmatrix} \cos\phi & \sin\phi & 0 & 0 \\ -\sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{C.60}$$

C.4 Inversion

The properties of a vector under inversion allows a classification of vectors:

$$\hat{\imath}\mathbf{r} = \begin{cases} -\mathbf{r} & \text{(polar) vector} \\ \mathbf{r} & \text{axial vector (pseuodvector)} \end{cases} \tag{C.61}$$

An example of a pseudo vector is the angular momentum vector. Correspondingly scalars can be classified as scalar or pseudo scalars.

$$\hat{\imath}a = \begin{cases} a & \text{scalar} \\ a & \text{pseudoscalar} \end{cases} \tag{C.62}$$

The inversion operator $\hat{\imath}$ commute with all rotations. This is straightforwardly seen from the effect of rotation and inversion on the coordinates:

$$\hat{R}(\phi,\mathbf{n})\,\hat{\imath}q_i = \hat{R}(\phi,\mathbf{n})(-q_i) = -q_jA_{ji}(\phi,\mathbf{n}) = \hat{\imath}\left(q_jA_{ji}(\phi,\mathbf{n})\right) = \hat{\imath}\hat{R}(\phi,\mathbf{n})\,q_i \tag{C.63}$$

From the above relation it becomes clear that the representation matrix for inversion in coordinate basis is

$$i_{\mathbf{r}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{C.64}$$

which is an ortogonal matrix with determinant -1. From the form of this matrix we may deduce the representation matrix for inversion in the basis of spin functions α and β:

$$\begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix} \otimes \begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{C.65}$$

We have set the scalar component equal to -1 corresponding to a pseudo scalar. The 4×4 inversion matrix is invariant under transformation to coordinate or spherical harmonic basis.

Note that the representation matrix for inversion in spin basis are of order four:

$$\mathbf{i} = -iI_2, \quad \mathbf{i}^2 = -I_2, \quad \mathbf{i}^3 = iI_2, \quad \mathbf{i}^4 = I_2 \tag{C.66}$$

C.5 Spatial symmetry in relativistic systems

The spin operation may be chosen freely since the non-relativistic Hamiltonian is spin free. We next consider a relativistic system and indicate this approxiamtely by adding the spin-orbnit operator $\hat{H}_{so}$

$$\hat{H}_{so} = c\boldsymbol{\sigma} \cdot (\nabla V \times \mathbf{p}) \tag{C.67}$$

We consider the transformation of $\hat{H}_{so}$ under a general symmetry operation $\hat{k}_i$(C.2):

$$\begin{aligned} & \hat{k}_r c\sigma_i \epsilon_{ijk} \tfrac{\delta V}{\delta q_j} p_k \hat{k}_r^\dagger \\ & = c\hat{G}^{\boldsymbol{\eta}}\left(\theta_\eta, \mathbf{m}_\eta, p'_\eta\right) \sigma_i \hat{G}^{\boldsymbol{\eta}}\left(-\theta_\eta, \mathbf{m}_\eta, p'_\eta\right) \epsilon_{ijk} \hat{G}\left(\phi_r, \mathbf{n_r}, p_r\right) \tfrac{\delta}{\delta q_j} \hat{G}\left(-\phi_r, \mathbf{n_r}, p_r\right) V \\ & \qquad\qquad \hat{G}\left(\phi_r, \mathbf{n_r}, p_r\right) p_k \hat{G}\left(-\phi_r, \mathbf{n_r}, p_r\right) \\ & = c\sigma_l \epsilon_{ijk} \tfrac{\delta V}{\delta q_m} p_n A_{li}\left(\theta, \mathbf{m}\right) A_{mj}\left(\phi_r, \mathbf{n_r}\right) A_{nk}\left(\phi_r, \mathbf{n_r}\right) \end{aligned} \tag{C.68}$$

where we have used that the potential V is totally symmetric under the point group. We see that $\hat{H}_{so}$ is not totally symmetric under $\hat{k}_i$. However, by inserting $\theta = \phi_r$ and $\mathbf{m} = \mathbf{n}_r$

we obtain

$$\begin{aligned}
&\hat{k}_r c\sigma_i \epsilon_{ijk} \tfrac{\delta V}{\delta q_j} p_k \hat{k}_r^\dagger \\
&= c\sigma_l \epsilon_{ijk} \tfrac{\delta V}{\delta q_m} p_n A_{li}\left(\phi_r, \mathbf{n_r}\right) A_{mj}\left(\phi_r, \mathbf{n_r}\right) A_{nk}\left(\phi_r, \mathbf{n_r}\right) \\
&= c\sigma_l \epsilon_{lmn} \tfrac{\delta V}{\delta q_m} p_n \det\left[\mathbf{A}\left(\phi_r, \mathbf{n}_r\right)\right] \\
&= c\sigma_l \epsilon_{lmn} \tfrac{\delta V}{\delta q_m} p_n
\end{aligned} \tag{C.69}$$

where we have used (C.10) and the fact that the rotation matrix $\mathbf{A}$ has determinant 1. From the above we see that symmetry operators in relativistic systems are generally given by:

$$\hat{k}_r = \hat{G}\left(\phi_r, \mathbf{n_r}, p_r\right) \hat{G}^{\boldsymbol{\eta}}\left(\phi_\eta, \mathbf{n_r}_\eta, p_\eta\right) \tag{C.70}$$

that is, the operation in spatial and spin coordinates must be identical.

C.6 Double groups

C.6.1 Binary symmetry operations

Let us first look at binary operations associated with the main axes In spin basis they are represented by

E	Identity	→	I_2
$C_2(x)$	Rotation π about the x-axis	→	$-i\sigma_x$
$C_2(y)$	Rotation π about the y-axis	→	$-i\sigma_y$
$C_2(z)$	Rotation π about the z-axis	→	$-i\sigma_z$
$\hat{i}$	Inversion	→	$-iI_2$
$\hat{\sigma}_y z$	Reflection in the yz-plane	→	$-\sigma_x$
$\hat{\sigma}_z x$	Reflection in the zx-plane	→	$-\sigma_y$
$\hat{\sigma}_x y$	Reflection in the xy plane	→	$-\sigma_z$

(C.71)

C.6.2 Example: D_2

Consider the group $D_2 = \{E, C_2(z), C_2(y), C_2(x)\}$. The group multiplication table gives

E	$C_2(z)$	$C_2(y)$	$C_2(x)$
$C_2(z)$	E	$C_2(x)$	$C_2(y)$
$C_2(y)$	$C_2(x)$	E	$C_2(z)$
$C_2(x)$	$C_2(y)$	$C_2(z)$	E

(C.72)

The corresponding multiplication table in spin representation is er:

$$
\begin{array}{r|rrr}
I_2 & -i\sigma_z & -i\sigma_y & -i\sigma_x \\
\hline
-i\sigma_z & -I_2 & -i\sigma_x & i\sigma_y \\
-i\sigma_y & i\sigma_x & -I_2 & -i\sigma_z \\
-i\sigma_x & -i\sigma_y & i\sigma_z & -I_2
\end{array}
\tag{C.73}
$$

By comparison we see that the multiplication tables differ by phase phactors

$$
\begin{array}{rrrr}
1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1
\end{array}
\tag{C.74}
$$

caused by the double periodicity of the elements. A correct representation may be obtained in two ways;

- by a projective (ray) representation instead of the regular vector representation

  $$G_iG_j = G_k \Rightarrow \mathbf{D}(G_i)\mathbf{D}(G_j) = \omega(G_i, G_j)\mathbf{D}(G_k) \tag{C.75}$$

  where $\omega(G_i, G_j)$ is phase factor that dependes of the order of operators G_i, G_j.
- Double groups: The periodicity of binary operations i extended from 2π to 4π.

Double groups are usually introduced by adding an extra element $\overline{E}$ representing a rotation 2π about an arbitrary axis and therefore commuting with all symmetry operations. We shall proceed a bit more stringently using our representation in spin basis The element $\overline{E}$ is the result of two binary operations about the same axis and is therefore represented by

$$\overline{E} = \hat{R}\,(\pi, \mathbf{n})\,\hat{R}\,(\pi, \mathbf{n}) \Rightarrow \overline{\mathbf{E}} = -i(\boldsymbol{\sigma}\cdot\mathbf{n}) - i(\boldsymbol{\sigma}\cdot\mathbf{n}) = -I_2 \tag{C.76}$$

By introducing the notation $\overline{C}_2 = C_2\overline{E}$ we obtain the following multiplication table

$$
\begin{array}{l|lllllll}
E & C_2(z) & C_2(y) & C_2(x) & \overline{E} & \overline{C}_2(z) & \overline{C}_2(y) & \overline{C}_2(x) \\
\hline
C_2(z) & \overline{E} & C_2(x) & \overline{C}_2(y) & \overline{C}_2(z) & E & C_2(x) & \overline{C}_2(y) \\
C_2(y) & C_2(x) & \overline{E} & \overline{C}_2(z) & \overline{C}_2(y) & \overline{C}_2(x) & E & C_2(z) \\
C_2(x) & \overline{C}_2(y) & C_2(z) & \overline{E} & \overline{C}_2(x) & C_2(y) & \overline{C}_2(z) & E \\
\overline{E} & \overline{C}_2(z) & \overline{C}_2(y) & \overline{C}_2(x) & E & C_2(z) & C_2(y) & C_2(x) \\
\overline{C}_2(z) & E & C_2(x) & \overline{C}_2(y) & C_2(z) & \overline{E} & \overline{C}_2(x) & C_2(y) \\
\overline{C}_2(y) & \overline{C}_2(x) & E & C_2(z) & C_2(y) & C_2(x) & \overline{E} & \overline{C}_2(z) \\
\overline{C}_2(x) & C_2(y) & \overline{C}_2(z) & E & C_2(x) & \overline{C}_2(y) & C_2(z) & \overline{E}
\end{array}
\tag{C.77}
$$

In order to construct a character table the distribution of symmetry operations among classes need to be determined. Two operations belong two the same class if there exists a third element in the group that bring the corresponding symmetry elements into each other. as we have introduced periodicity 4π , C_2 og C_2^{-1} are no longer identical. C_2 og C_2^{-1} belong to the same class if the group contains a binary rotation normal to the rotation axis or if the group contains a vertical plane. A horizontal plane will reverse the direction of the axis, but also the direction of rotation, so that the two effects cancel out.

In the double group $\overline{D}_2$ we have binary operations perpendicular two each other, so that elements C_2 og $\overline{C_2}$ are in the same class. The double group $\overline{D}_2$ therefore has five classes $\{(E), (\overline{E}), (C_2(z), \overline{C}_2(z)), (C_2(y), \overline{C}_2(y)), (C_2(x), \overline{C}_2(x))\}$, and thereby five irreducible representations. We have one extra irrep compared to the single group D_2. The extra irrep is spanned by functions with half-integer values of angular momentum and they are therefore denoted fermion irreps. The regular irreps spanned by integer values of angular momentum are the termed boson irreps.

The double group $\overline{D}_2$ has a total of eight elements, so that the extra irrep must be twodimensional (it follows from the conditiona that the sum of squares of the dimension of irreps equals the number of operations). This allows us to construct the following character table:

$\overline{D}_2$	E	$\overline{E}$	$2C_2(z)$	$2C_2(y)$	$2C_2(x)$		
A	1	1	1	1	1		$\frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)$
B_1	1	1	1	-1	-1	z	$-\frac{1}{\sqrt{2}}(\alpha\beta + \beta\alpha)$
B_2	1	1	-1	1	-1	y	$\frac{i}{\sqrt{2}}(\alpha\alpha + \beta\beta)$
B_3	1	1	-1	-1	1	x	$\frac{1}{\sqrt{2}}(\alpha\alpha - \beta\beta)$
$E_{\frac{1}{2}}$	2	-2	0	0	0		(α, β)

(C.78)

We see that the boson irreps duplicate the correponding irreps of the single groups, which is understandable since they are spanned by functions for which the symmetry operations have periodicity 2π. The character of E for the fermion irrep follows from its dimensionality. The same holds for the character of $\overline{E}$, but now with a minus sign since $\overline{E} = -I_{2j+1}$. The character of the other operations can be deduced from the little orthogonality theorem.

In the character table we give examples of spin functions that span the various irreps. They can be obtained by projection operator. Note that the components of the triplet do not span separate boson irreps, but form linear combinations corresponding to the characters.

Appendix D

Diagonalization of quaternion Hermitian matrix

We will consider the diagonalization of a quaternion Hermitian matrix

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1\check{\imath} + \mathbf{H}_2\check{\jmath} + \mathbf{H}_3\check{\mathrm{k}} = \mathbf{H}_0^T - \mathbf{H}_1^T\check{\imath} - \mathbf{H}_2^T\check{\jmath} - \mathbf{H}_3^T\check{\mathrm{k}} = \mathbf{H}^\dagger \tag{D.1}$$

From the condition of Hermiticity we see that $\mathbf{H}_0$ is a real symmetric matrix, whereas $\mathbf{H}_{\{1-3\}}$ are real antisymmetric matrices. The diagonalization of $\mathbf{H}$ proceeds in four steps:

1. Reduction to quaternion Hermitian tridiagonal matrix by a quaternion analogue of the Householder method
2. Reduction to real symmetric tridiagonal matrix
3. Diagonalization of real symmetric tridiagonal matrix
4. Backtransformation to obtain eigenvectors

The first two steps are handled by the routine `QHTRID` and the final step by `QHTRBK`. The third step is hadled by the standard `TQL2` routine using the QL algorithm which is described in [169]. When only eigenvalues and no eigenvectors are desired, the routine `TQLRAT` is called instead of `TQL2` and execution terminated after obtaining the eigenvalues.

A quaternionic number q and its conjugate q^* is given by

$$q = a + b\check{\imath} + c\check{\jmath} + d\check{\mathrm{k}}; \qquad q^* = a - b\check{\imath} - c\check{\jmath} - d\check{\mathrm{k}} \tag{D.2}$$

in which the quaternion units $\check{\imath}$, $\check{\jmath}$ and $\check{\mathrm{k}}$ obey the following multiplication rules

$$\check{\imath}^2 = \check{\jmath}^2 = \check{\mathrm{k}}^2 = \check{\imath}\check{\jmath}\check{\mathrm{k}} = -1 \tag{D.3}$$

From these multiplication rules it follows that quaternion number do not commute under multiplication

$$\begin{aligned} q_a q_b &= [a_1 + ia_2 + ja_3 + ka_4]\,[b_1 + ib_2 + jb_3 + kb_4] \\ &= \quad [a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4] \\ &+ \quad \check{\imath}\,[a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3] \\ &+ \quad \check{\jmath}\,[a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2] \\ &+ \quad \check{k}\,[a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1] \quad \neq \quad q_b q_a \end{aligned} \tag{D.4}$$

Note, however, that the real part of the product is independent of order. This means that in devising an algorithm for quaternion diagonalization great care has to be taken in not reversing the order of quaternion multiplication. The routine for quaternion number is in its entirely written in terms of real variables.

D.1 Quaternion Householder matrices

We consider a quaternion Householder matrix $\mathbf{P}$ of the form

$$\mathbf{P} = \mathbf{1} - 2\mathbf{w}\cdot\mathbf{w}^\dagger; \qquad |\mathbf{w}|^2 = \mathbf{w}^\dagger\mathbf{w} = 1 \tag{D.5}$$

The matrix $\mathbf{P}$ is unitary as

$$\begin{aligned} \mathbf{P}^\dagger\mathbf{P} &= \mathbf{P}^2 = \left[\mathbf{1} - 2\mathbf{w}\cdot\mathbf{w}^\dagger\right]\left[\mathbf{1} - 2\mathbf{w}\cdot\mathbf{w}^\dagger\right] \\ &= \mathbf{1} - 2\mathbf{w}\cdot\mathbf{w}^\dagger - 2\mathbf{w}\cdot\mathbf{w}^\dagger + 4\left[\mathbf{w}\cdot\mathbf{w}^\dagger\right]\left[\mathbf{w}\cdot\mathbf{w}^\dagger\right] \\ &= \mathbf{1} - 2\mathbf{w}\cdot\mathbf{w}^\dagger - 2\mathbf{w}\cdot\mathbf{w}^\dagger + 4\mathbf{w}\,|\mathbf{w}|^2\cdot\mathbf{w}^\dagger \qquad = \quad \mathbf{1} \end{aligned} \tag{D.6}$$

We now express the matrix $\mathbf{P}$ as

$$\mathbf{P} = \mathbf{1} - \frac{\mathbf{u}\cdot\mathbf{u}^\dagger}{H}; \qquad H \equiv \frac{1}{2}\,|\mathbf{u}|^2 \tag{D.7}$$

where $\mathbf{u}$ can be any vector. We choose

$$\mathbf{u} = \mathbf{x} + |\mathbf{x}|\,\frac{x_1}{|x_1|}\mathbf{e}_1 \tag{D.8}$$

where $\mathbf{e}_1$ is the unit vector $[1, 0, \ldots, 0]^T$ and $\mathbf{x}$ is an arbitrary quaternion vector. This gives

$$\begin{aligned} H &= \tfrac{1}{2}\mathbf{u}^\dagger\mathbf{u} \\ &= \left[\mathbf{x}^\dagger + |\mathbf{x}|\,\frac{x_1^*}{|x_1|}\mathbf{e}_1\right]\cdot\left[\mathbf{x} + |\mathbf{x}|\,\frac{x_1}{|x_1|}\mathbf{e}_1\right] \\ &= |\mathbf{x}|^2 + |\mathbf{x}|\,|x_1| \end{aligned} \tag{D.9}$$

We then consider the action of $\mathbf{P}$ on the vector $\mathbf{x}$:

$$\begin{aligned} \mathbf{P}\cdot\mathbf{x} &= \mathbf{x} - \frac{\mathbf{u}}{H}\cdot\left(\mathbf{x}^\dagger + |\mathbf{x}|\,\tfrac{x_1^*}{|x_1|}\mathbf{e}_1\right)\cdot\mathbf{x} \\ &= \mathbf{x} - \frac{\mathbf{u}\cdot\left[|\mathbf{x}|^2 + |\mathbf{x}|\,\tfrac{|x_1|^2}{|x_1|}\right]}{|\mathbf{x}|^2 + |\mathbf{x}|\,|x_1|} \\ &= \mathbf{x} - \mathbf{u} \\ &= |\mathbf{x}|\,\tfrac{x_1}{|x_1|}\mathbf{e}_1 \end{aligned} \tag{D.10}$$

We see that $\mathbf{P}$ operating on $\mathbf{x}$ gives a vector in which all elements are zero except for the first element, which is $\frac{x_1}{|x_1|}$ We may note that operating on $\mathbf{x}$ with

$$\frac{x_1^*}{|x_1|}\mathbf{P} \tag{D.11}$$

gives a real vector in which all elements are zero except for the first , which is $|\mathbf{x}|$.

D.2 Reduction to quaternion Hermitian tridiagonal matrix

We want to reduce a quaternion $n \times n$ Hermitian matrix $\mathbf{H}$ to a quaternion Hermitian tridiagonal matrix $\mathbf{T}$ through a finite series of quaternion unitary transformations. Using quaternion Householder matrices (D.7) this may be accomplished in $n-2$ steps.

Our quaternion Hermitian matrix $\mathbf{H}$ may be written as

$$\mathbf{H} = \mathbf{H}^{(0)} = \begin{array}{c|cc} & [n-1] & [1] \\ \hline [n-1] & \mathbf{A}^{(0)} & \mathbf{b}^{(0)} \\ [1] & \mathbf{b}^{(0)\dagger} & \mathbf{B}^{(0)} \end{array} \tag{D.12}$$

where

$$\mathbf{b}^{(0)\dagger} = [h_{n1}, h_{n2}, h_{n3}, \ldots, h_{n,n-1}]\,; \qquad \mathbf{B}^{(0)} = h_{nn} \tag{D.13}$$

We now choose a Householder matrix of the form

$$\mathbf{P}^{(1)} = \begin{bmatrix} \mathbf{Q}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_1 \end{bmatrix} : \qquad \mathbf{Q}^{(1)} = \mathbf{I}_{(n-1)} - \frac{\mathbf{u}^{(1)}\mathbf{u}^{(1)\dagger}}{H}; \quad H = \frac{1}{2}\left|\mathbf{u}^{(1)}\right|^2 \tag{D.14}$$

with

$$\mathbf{u}^{(1)\dagger} = \mathbf{b}^{(0)\dagger} + \left|\mathbf{b}^{(0)}\right| \frac{b_{(n-1)}^{(0)*}}{\left|b_{(n-1)}^{(0)}\right|}\mathbf{e}_{(n-1)}; \quad H = \left|\mathbf{b}^{(0)}\right|^2 + \left|\mathbf{b}^{(0)}\right| \left|b_{(n-1)}^{(0)}\right| \tag{D.15}$$

The unitary transformation $\mathbf{P}^{(1)}\mathbf{A}^{(0)}\mathbf{P}^{(1)}$ then gives

$$\mathbf{P}^{(1)}\mathbf{A}^{(0)}\mathbf{P}^{(1)} = \begin{bmatrix} \mathbf{Q}^{(1)}\mathbf{A}^{(0)}\mathbf{Q}^{(1)} & \mathbf{Q}^{(1)}\mathbf{b}^{(0)} \\ \mathbf{b}^{(0)\dagger}\mathbf{Q}^{(1)} & \mathbf{B}^{(0)} \end{bmatrix} \tag{D.16}$$

with

$$\mathbf{b}^{(0)}\mathbf{Q}^{(1)} = \left|\mathbf{b}^{(0)}\right| \frac{b^{(0)*}_{(n-1)}}{\left|b^{(0)}_{(n-1)}\right|}\mathbf{e}_{(n-1)} = \left[0, 0, \ldots, 0, \left|\mathbf{b}^{(0)}\right| \frac{h^{(0)}_{(n-1)}}{\left|h^{(0)}_{(n-1)}\right|}\right] \tag{D.17}$$

We see that the elements of $\mathbf{b}^{(0)}$ has been zeroed out, except for the final element. We may now write the transformed matrix as:

$$\mathbf{H}^{(1)} = \begin{array}{c|cc} & [n-2] & [2] \\ \hline [n-2] & \mathbf{A}^{(1)} & \begin{bmatrix} \mathbf{b}^{(1)} & 0 \end{bmatrix} \\ [2] & \begin{bmatrix} \mathbf{b}^{(1)\dagger} \\ 0 \end{bmatrix} & \mathbf{B}^{(1)} \end{array} \tag{D.18}$$

where

$$\mathbf{B}^{(1)} = \begin{bmatrix} h^{(1)}_{(n-1),(n-1)} & \left|\mathbf{b}^{(0)}\right| \frac{b^{(0)}_{(n-1)}}{\left|b^{(0)}_{(n-1)}\right|} \\ \left|\mathbf{b}^{(0)}\right| \frac{b^{(0)*}_{(n-1)}}{\left|b^{(0)}_{(n-1)}\right|} & h^{(0)}_{nn} \end{bmatrix} \tag{D.19}$$

After step $(i-1)$ our quaternion matrix has the structure

$$\mathbf{H}^{(i-1)} = \begin{array}{c|cc} & [n-i] & [i] \\ \hline [n-i] & \mathbf{A}^{(i-1)} & \left[\mathbf{b}^{(i-1)} \quad 0^{(i-1)} \right) \\ [i] & \begin{bmatrix} \mathbf{b}^{(1)\dagger} \\ 0^{(i-1)\dagger} \end{bmatrix} & \mathbf{B}^{(i-1)} \end{array} \tag{D.20}$$

Here $0^{(i-1)}$ is a $(i-1) \times (n-i)$ zero matrix while $\mathbf{B}^{(i-1)}$ is a $i \times i$ quaternion tridiagonal matrix. For step i we introduce the index m defined by $m = n - i$ and choose the Householder matrix

$$\mathbf{P}^{(i)} = \begin{bmatrix} \mathbf{Q}^{(i)} & 0_i \\ 0_i & \mathbf{I}_i \end{bmatrix}; \qquad \mathbf{Q}^{(i)} = \mathbf{I}_m - \frac{\mathbf{u}^{(i)}\mathbf{u}^{(i)\dagger}}{H}; \qquad H = \frac{1}{2}\left|\mathbf{u}^{(\mathbf{i})}\right|^2 \tag{D.21}$$

with

$$\mathbf{u}^{(i)\dagger} = \mathbf{b}^{(i-1)\dagger} + \left|\mathbf{b}^{(i-1)}\right| \frac{b^{(i-1)*}_m}{\left|b^{(i)}_m\right|}\mathbf{e}_m; \quad H = \left|\mathbf{b}^{(i)}\right|^2 + \left|\mathbf{b}^{(i)}\right| \left|b^{(i)}_m\right| \tag{D.22}$$

After $(n-2)$ steps the quaternion Hermitian matrix $\mathbf{H}$ has been reduced to a quaternion Hermitian tridiagonal matrix $\mathbf{T}$:

$$\mathbf{T} = \mathbf{V}\mathbf{H}\mathbf{V}^{\dagger}; \quad \mathbf{V} = \prod_{i=1}^{n-2} \mathbf{P}^{(i)} \tag{D.23}$$

D.3 Reduction to real symmetric tridiagonal matrix

The quaternion Hermitian tridiagonal matrix $\mathbf{T}_Q$ has the structure

$$\mathbf{T}_Q = \begin{bmatrix} h_{11}^{(n-2)} & h_{21}^{(n-2)*} & 0 & \dots & 0 & 0 & 0 \\ h_{21}^{(n-2)} & h_{22}^{(n-2)} & c^{(n-2)*} & \dots & 0 & 0 & 0 \\ 0 & c^{(n-2)} & h_{33}^{(n-3)} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & h_{(n-2),(n-2)}^{(1)} & c^{(2)*} & 0 \\ 0 & 0 & 0 & \dots & c^{(2)} & h_{(n-1),(n-1)}^{(1)} & c^{(1)*} \\ 0 & 0 & 0 & \dots & 0 & c^{(1)} & h_{nn}^{(0)} \end{bmatrix} \tag{D.24}$$

where

$$c^{(i)} = \left|b^{(i-1)}\right| \frac{b_{(m)}^{(i-1)*}}{\left|b_{(m)}^{(i-1)}\right|} = \left|b^{(i-1)}\right| \frac{h_{(m+1),m}^{(i-1)}}{\left|h_{(m+1),m}^{(i-1)}\right|} \tag{D.25}$$

We now consider the unitary transformation

$$\mathbf{T}_R = \boldsymbol{\tau}^{\dagger}\mathbf{T}_Q\boldsymbol{\tau}; \qquad \boldsymbol{\tau}^{\dagger}\boldsymbol{\tau} = \mathbf{I}; \quad \tau_{ij} = \tau_i\delta_{ij} \tag{D.26}$$

transforming $\mathbf{T}_Q$ to a real symmetric tridiagonal matrix $\mathbf{T}_R$ using the quaternion unitary diagonal matrix $\boldsymbol{\tau}$. For the case $n = 3$ the transformation is

$$\begin{aligned} \mathbf{T}_R = \boldsymbol{\tau}^{\dagger}\mathbf{T}_Q\boldsymbol{\tau} &= \begin{bmatrix} \tau_1^* & 0 & 0 \\ 0 & \tau_2^* & 0 \\ 0 & 0 & \tau_3^* \end{bmatrix} \begin{bmatrix} t_{11} & t_{21}^* & 0 \\ t_{21} & t_{22} & t_{32}^* \\ 0 & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} \tau_1 & 0 & 0 \\ 0 & \tau_2 & 0 \\ 0 & 0 & \tau_3 \end{bmatrix} \\ &= \begin{bmatrix} \tau_1^* t_{11}\tau_1 & \tau_1^* t_{21}^*\tau_2 & 0 \\ \tau_2^* t_{21}\tau_1 & \tau_2^* t_{22}\tau_2 & \tau_2^* t_{32}^*\tau_3 \\ 0 & \tau_3^* t_{32}\tau_2 & \tau_3^* t_{33}\tau_3 \end{bmatrix} \end{aligned} \tag{D.27}$$

Using the unitarity of $\boldsymbol{\tau}$ and the fact that the diagonal elements of $\mathbf{T}_Q$ are real, we obtain:

$$\mathbf{T}_R = \boldsymbol{\tau}^\dagger \mathbf{T}_Q \boldsymbol{\tau} = \begin{bmatrix} t_{11} & \tau_1^* t_{21}^* \tau_2 & 0 \\ \tau_2^* t_{21} \tau_1 & t_{22} & \tau_2^* t_{32}^* \tau_3 \\ 0 & \tau_3^* t_{32} \tau_2 & t_{33} \end{bmatrix} \tag{D.28}$$

We then find:

$$\begin{aligned} \tau_3 &= 1 \\ \tau_2 &= \frac{t_{32}^*}{|t_{32}|} \\ \tau_1 &= \frac{t_{21}^*}{|t_{21}|}\frac{t_{32}^*}{|t_{32}|} = \frac{t_{21}^*}{|t_{21}|}\tau_2 \end{aligned} \tag{D.29}$$

Generalizing we find:

$$\begin{aligned} \tau_m &= \frac{c^{(i)*}}{\left|c^{(i)}\right|}\tau_{(m+1)} = \frac{h^{(i-1)*}_{(m+1),m}}{\left|h^{(i-1)}_{(m+1),m}\right|}\tau_{(m+1)} \qquad (1 < m < n) \\ \tau_n &= 1 \end{aligned} \tag{D.30}$$

D.4 Implementation of tridiagonalization

We now consider the computational expressions for the reduction of a quaternion Hermitian matrix to a real symmetric tridiagonal matrix. We first consider the matrix:

$$\mathbf{A}^{(i-1)}\mathbf{Q}^{(i)} = \mathbf{A}^{(i-1)}\left[\mathbf{I}_m - \frac{\mathbf{u}^{(i)}\mathbf{u}^{(i)\dagger}}{H}\right] = \mathbf{A}^{(i-1)} - \mathbf{p}^{(i)}\mathbf{u}^{(i)\dagger} \tag{D.31}$$

where we have introduced the vector $\mathbf{p}^{(i)}$:

$$\mathbf{p}^{(i)} \equiv \frac{\mathbf{A}^{(i-1)}\mathbf{u}^{(i)}}{H} \tag{D.32}$$

Note that due to the hermiticity of $\mathbf{A}$ we have:

$$\mathbf{p}^{(i)\dagger} = \frac{\mathbf{u}^{(i)\dagger}\mathbf{A}^{(i-1)}}{H} \tag{D.33}$$

The full Householder transformation of the subblock $\mathbf{A}^{(i-1)}$ of the matrix $\mathbf{H}^{(i-1)}$ then becomes:

$$\begin{aligned} \mathbf{Q}^{(i)}\mathbf{A}^{(i-1)}\mathbf{Q}^{(i)} &= \left[\mathbf{I}_m - \frac{\mathbf{u}^{(i)}\mathbf{u}^{(i)\dagger}}{H}\right]\left[\mathbf{A}^{(i-1)} - \mathbf{p}^{(i)}\mathbf{u}^{(i)\dagger}\right] \\ &= \mathbf{A}^{(i-1)} - \mathbf{p}^{(i)}\mathbf{u}^{(i)\dagger} - \mathbf{u}^{(i)}\mathbf{p}^{(i)\dagger} + 2\mathbf{u}^{(i)}K\mathbf{u}^{(i)\dagger} \end{aligned} \tag{D.34}$$

where we have introduced the scalar K defined by

$$K \equiv \left[\frac{\mathbf{u}^{(i)\dagger}\mathbf{A}^{(i-1)}\mathbf{u}}{2H^2} \right] \tag{D.35}$$

K is real due to the hermiticity of **A**:

$$K^* = \left[\frac{\mathbf{u}^{(i)\dagger}\mathbf{A}^{(i-1)\dagger}\mathbf{u}}{2H^2} \right] = K \tag{D.36}$$

If we write

$$\mathbf{q} \equiv \mathbf{p} - K\mathbf{u} \tag{D.37}$$

then we have

$$\mathbf{Q}^{(i)}\mathbf{A}^{(i-1)}\mathbf{Q}^{(i)} = \mathbf{A}^{(i-1)} - \mathbf{q}^{(i)}\mathbf{u}^{(i)\dagger} - \mathbf{u}^{(i)}\mathbf{q}^{(i)\dagger} \tag{D.38}$$

This is the computationally useful formula.

D.5 Eigenvectors

The complete tranformation of the quaternion Hermitian matrix **H** to a real diagonal matrix **Λ** is given by

$$\Lambda = \mathbf{U}^\dagger\mathbf{H}\mathbf{U}; \qquad \mathbf{U} = \mathbf{V}\boldsymbol{\tau}\mathbf{O} \tag{D.39}$$

with

$$\begin{array}{lcll} \mathbf{V} & = & \prod_{i=1}^{n-2}\mathbf{P}^{(i)} & \text{- reduction to quaternion Hermitian tridiagonal matrix} \\ \boldsymbol{\tau} & & & \text{-reduction to real symmetric tridiagonal matrix} \\ \mathbf{O} & & & \text{-reduction to real diagonal matrix} \end{array} \tag{D.40}$$

The transformation from a real tridiagonal to real diagonal matrix is handled by the QL algorithm, described in [169]. From the eigenvectors obtained from the real tridiagonal matrix we may backtransform to the eigenvectors of the quaternion Hermitian matrix. In the first step the eigenvectors of the quaternion Hermitian tridiagonal matrix are recovered:

$$\mathbf{U}^{(2)} = [\boldsymbol{\tau}O]_{kj} = \sum_{i=1}^{n} \tau_{ik}\delta_{ik}o_{kj} = \tau_{kk}o_{kj} \tag{D.41}$$

Then follows (n-2) steps in which the eigenvectors of the full quaternion Hermitian matrix are recovered. In step i we have

$$\mathbf{Z}^{(i)} = \mathbf{P}^{(i)}\mathbf{Z}^{(i-1)} = \left[\begin{array}{cc} \mathbf{Q}_{mm} & \mathbf{0}_{mi} \\ \mathbf{0}_{im} & \mathbf{I}_{ii} \end{array}\right] \left[\begin{array}{cc} \mathbf{Z}_{mm} & \mathbf{Z}_{mi} \\ \mathbf{Z}_{im} & \mathbf{Z}_{ii} \end{array}\right] = \left[\begin{array}{c} \mathbf{Q}_{mm}\mathbf{Z}_{mn} \\ \mathbf{Q}_{im}\mathbf{Z}_{in} \end{array}\right] \tag{D.42}$$

A computational useful formula is found by expansion:

$$[\mathbf{QZ}]_{kj} = \sum_l \mathbf{Q}_{kl}\mathbf{Z}_{lj} = \sum_l \left(\delta_{kl} - \mathbf{u}_k \frac{\mathbf{u}_l^\dagger}{H}\right) \mathbf{Z}_{lj} = \mathbf{Z}_{kj} - \mathbf{u}_k \mathbf{s}_j \tag{D.43}$$

where we have introduced

$$\mathbf{s}_j = \sum_l \frac{\mathbf{u}_l^\dagger}{H} \mathbf{Z}_{lj} \tag{D.44}$$

Appendix E

Angular part of atomic 2-spinors

In these notes angular 2-spinors and corresponding spherical harmonic functions are tabulated and plotted. The density plots appearing in these notes have been made by Jon K. Lærdahl.

The angular functions χ_{κ,m_j} appearing in the solutions of the Dirac equation for hydrogenlike atoms can be written the angular part of the hydrogenic solutions to the Dirac equation

$$\chi_{\kappa,m_j} = \frac{1}{\sqrt{2l+1}} \begin{bmatrix} a\sqrt{l+\frac{1}{2}+am_j}Y_l^{m_j-\frac{1}{2}} \\ \\ \sqrt{l+\frac{1}{2}-am_j}Y_l^{m_j+\frac{1}{2}} \end{bmatrix} \tag{E.1}$$

where

$$\kappa = a(j+1/2); \qquad a = 2(j-l) = \pm 1 \tag{E.2}$$

The Y_l^m are spherical harmonic functions (with the Condon-Shortley phase convention) [70]

$$Y_l^m(\theta,\phi) \equiv (-1)^m \sqrt{\frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\phi)\, e^{im\phi} \tag{E.3}$$

defined in terms of associated Legendre functions

$$P_l^m(x) = \frac{1}{2^l l!}\left(1-x^2\right)^{m/2} \frac{d^{l+m}}{dx^{l+m}}\left(x^2-1\right)^l; \qquad -l \le m \le l \tag{E.4}$$

E.1 Spherical harmonics

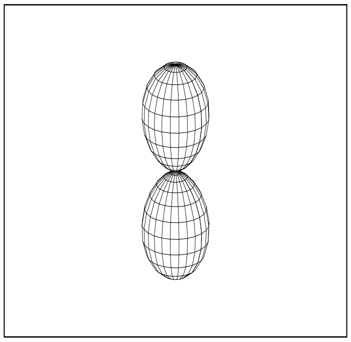
For reference we tabulate the spherical harmonics up to l = 3

l	m_l	$N_{l,m}$	polar form	Cartesian form
1	0	$\sqrt{\frac{3}{4\pi}}$	$\cos\Theta$	z
1	±1	$\mp\sqrt{\frac{3}{8\pi}}$	$\sin\Theta e^{\pm i\phi}$	$(x \pm iy)$
2	0	$\sqrt{\frac{5}{16\pi}}$	$\left(3\cos^2\Theta - 1\right)$	$2z^2 - x^2 - y^2$
2	±1	$\mp\sqrt{\frac{15}{8\pi}}$	$\cos\Theta\sin\Theta e^{\pm i\phi}$	$xz \pm iyz$
2	±2	$\sqrt{\frac{15}{32\pi}}$	$\sin^2\Theta e^{\pm 2i\phi}$	$x^2 \pm 2ixy - y^2$
3	0	$\sqrt{\frac{63}{16\pi}}$	$\left(\frac{5}{3}\cos^3\Theta - \cos\Theta\right)$	$2z^3 - 3x^2z - 3y^2z$
3	±1	$\mp\sqrt{\frac{21}{64\pi}}$	$\sin\Theta\left(5\cos^2\Theta - 1\right)e^{\pm i\phi}$	$4xz^2 \pm i4yz^2 - x^3 \mp ix^2y - xy^2 \mp iy^3$
3	±2	$\sqrt{\frac{105}{32\pi}}$	$\sin^2\Theta\cos\Theta e^{\pm 2i\phi}$	$x^2z \pm 2ixyz - y^2z$
3	±3	$\mp\sqrt{\frac{35}{64\pi}}$	$\sin^3\Theta e^{\pm 3i\phi}$	$x^3 \pm 3ix^2y - 3xy^2 \mp iy^3$

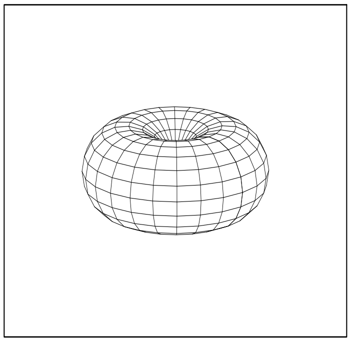
(E.5)

Density Plots:

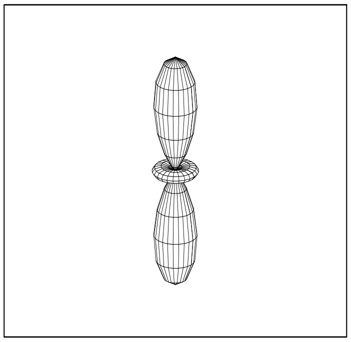
Y_{10} :



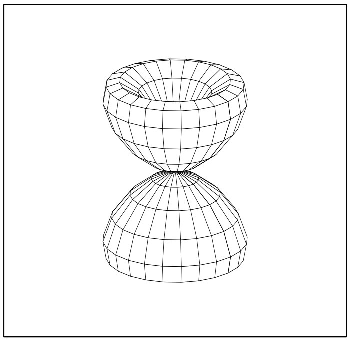
$Y_{1\pm1}$:



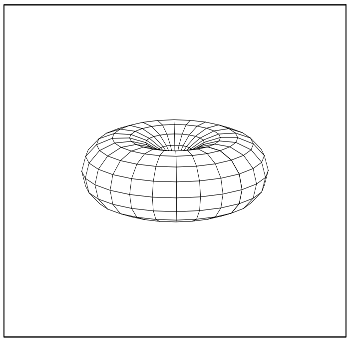
$Y_{20}:$



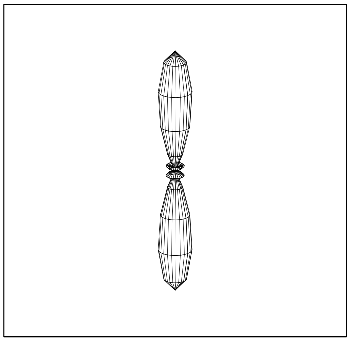
$Y_{2\pm1}:$



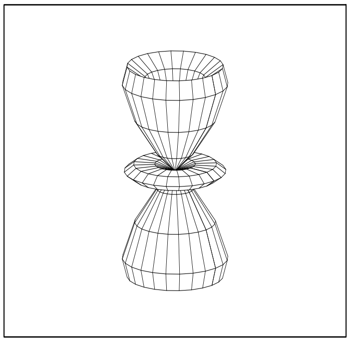
$Y_{2\pm2}:$



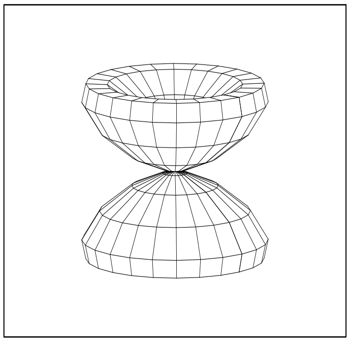
$Y_{30}:$



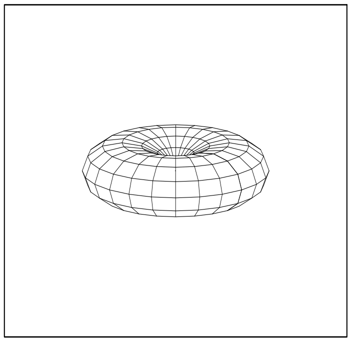
$Y_{3\pm1}:$



$Y_{3\pm2}:$



$Y_{3\pm3}:$



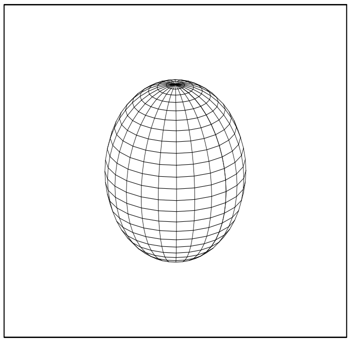
E.2 2-spinors

In this section we list the angular 2-spinors up to $j = 7/2$. We use the notation $(l)(j, m_j)$ where $(l) = s, p, d, f$ with reference to the angular quantum number l.

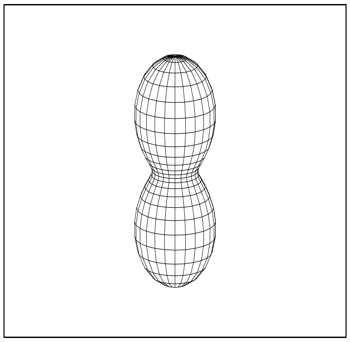
orb	α	β	*orb*	α	β
			$p\left(\frac{3}{2},\frac{3}{2}\right)$	$Y_{1,1}$	
$p\left(\frac{1}{2},\frac{1}{2}\right)$	$\sqrt{\frac{1}{3}}Y_{1,0}$	$-\sqrt{\frac{2}{3}}Y_{1,1}$	$p\left(\frac{3}{2},\frac{1}{2}\right)$	$\sqrt{\frac{2}{3}}Y_{1,0}$	$\sqrt{\frac{1}{3}}Y_{1,1}$
$p\left(\frac{1}{2},-\frac{1}{2}\right)$	$\sqrt{\frac{2}{3}}Y_{1,-1}$	$-\sqrt{\frac{1}{3}}Y_{1,0}$	$p\left(\frac{3}{2},-\frac{1}{2}\right)$	$\sqrt{\frac{1}{3}}Y_{1,-1}$	$\sqrt{\frac{2}{3}}Y_{1,0}$
			$p\left(\frac{3}{2},-\frac{3}{2}\right)$		$Y_{1,-1}$
			$d\left(\frac{5}{2},\frac{5}{2}\right)$	$Y_{2,2}$	
$d\left(\frac{3}{2},\frac{3}{2}\right)$	$\sqrt{\frac{1}{5}}Y_{2,1}$	$-\sqrt{\frac{4}{5}}Y_{2,2}$	$d\left(\frac{5}{2},\frac{3}{2}\right)$	$\sqrt{\frac{4}{5}}Y_{2,1}$	$\sqrt{\frac{1}{5}}Y_{2,2}$
$d\left(\frac{3}{2},\frac{1}{2}\right)$	$\sqrt{\frac{2}{5}}Y_{2,0}$	$-\sqrt{\frac{3}{5}}Y_{2,1}$	$d\left(\frac{5}{2},\frac{1}{2}\right)$	$\sqrt{\frac{3}{5}}Y_{2,0}$	$\sqrt{\frac{2}{5}}Y_{2,1}$
$d\left(\frac{3}{2},-\frac{1}{2}\right)$	$\sqrt{\frac{3}{5}}Y_{2,-1}$	$-\sqrt{\frac{2}{5}}Y_{2,0}$	$d\left(\frac{5}{2},-\frac{1}{2}\right)$	$\sqrt{\frac{2}{5}}Y_{2,-1}$	$\sqrt{\frac{3}{5}}Y_{2,0}$
$d\left(\frac{3}{2},-\frac{3}{2}\right)$	$\sqrt{\frac{4}{5}}Y_{2,-2}$	$-\sqrt{\frac{1}{5}}Y_{2,-1}$	$d\left(\frac{5}{2},-\frac{3}{2}\right)$	$\sqrt{\frac{1}{5}}Y_{2,-2}$	$\sqrt{\frac{4}{5}}Y_{2,-1}$
			$d\left(\frac{5}{2},-\frac{5}{2}\right)$		$Y_{2,-2}$
			$f\left(\frac{7}{2},\frac{7}{2}\right)$	$Y_{3,3}$	
$f\left(\frac{5}{2},\frac{5}{2}\right)$	$\sqrt{\frac{1}{7}}Y_{3,2}$	$-\sqrt{\frac{6}{7}}Y_{3,3}$	$f\left(\frac{7}{2},\frac{5}{2}\right)$	$\sqrt{\frac{6}{7}}Y_{3,2}$	$\sqrt{\frac{1}{7}}Y_{3,3}$
$f\left(\frac{5}{2},\frac{3}{2}\right)$	$\sqrt{\frac{2}{7}}Y_{3,1}$	$-\sqrt{\frac{5}{7}}Y_{3,2}$	$f\left(\frac{7}{2},\frac{3}{2}\right)$	$\sqrt{\frac{5}{7}}Y_{3,1}$	$\sqrt{\frac{2}{7}}Y_{3,2}$
$f\left(\frac{5}{2},\frac{1}{2}\right)$	$\sqrt{\frac{3}{7}}Y_{3,0}$	$-\sqrt{\frac{4}{7}}Y_{3,1}$	$f\left(\frac{7}{2},\frac{1}{2}\right)$	$\sqrt{\frac{4}{7}}Y_{3,0}$	$\sqrt{\frac{3}{7}}Y_{3,1}$
$f\left(\frac{5}{2},-\frac{1}{2}\right)$	$\sqrt{\frac{4}{7}}Y_{3,-1}$	$-\sqrt{\frac{3}{7}}Y_{3,0}$	$f\left(\frac{7}{2},-\frac{1}{2}\right)$	$\sqrt{\frac{3}{7}}Y_{3,-1}$	$\sqrt{\frac{4}{7}}Y_{3,0}$
$f\left(\frac{5}{2},-\frac{3}{2}\right)$	$\sqrt{\frac{5}{7}}Y_{3,-2}$	$-\sqrt{\frac{2}{7}}Y_{3,-1}$	$f\left(\frac{7}{2},-\frac{3}{2}\right)$	$\sqrt{\frac{2}{7}}Y_{3,-2}$	$\sqrt{\frac{5}{7}}Y_{3,-1}$
$f\left(\frac{5}{2},-\frac{5}{2}\right)$	$\sqrt{\frac{6}{7}}Y_{3,-3}$	$-\sqrt{\frac{1}{7}}Y_{3,-2}$	$f\left(\frac{7}{2},-\frac{5}{2}\right)$	$\sqrt{\frac{1}{7}}Y_{3,-3}$	$\sqrt{\frac{6}{7}}Y_{3,-2}$
			$f\left(\frac{7}{2},-\frac{7}{2}\right)$		$Y_{3,-3}$

Density Plots:

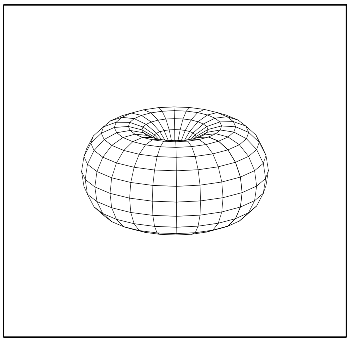
$s\left(\frac{1}{2},\frac{1}{2}\right), p\left(\frac{1}{2},\frac{1}{2}\right):$



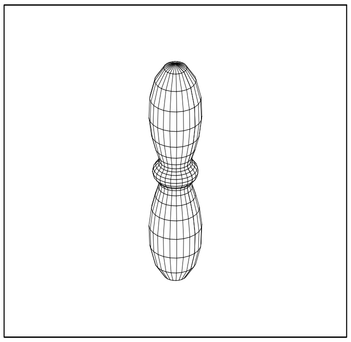
$p\left(\frac{3}{2},\frac{1}{2}\right), d\left(\frac{3}{2},\frac{1}{2}\right):$



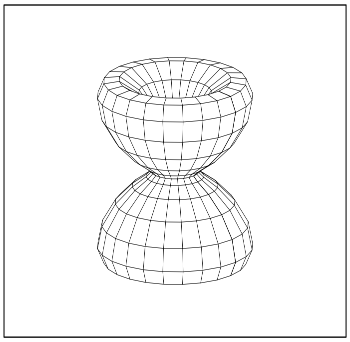
$p\left(\frac{3}{2},\frac{3}{2}\right), d\left(\frac{3}{2},\frac{3}{2}\right):$



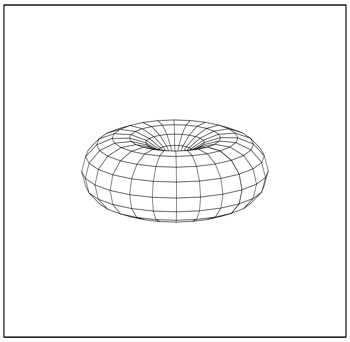
$d\left(\frac{5}{2},\frac{1}{2}\right), f\left(\frac{5}{2},\frac{1}{2}\right):$



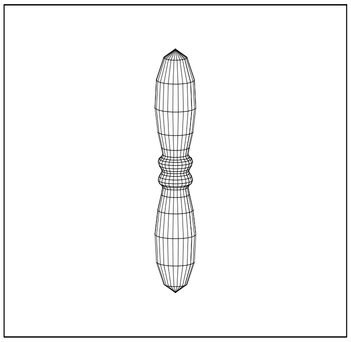
$d\left(\frac{5}{2},\frac{3}{2}\right), f\left(\frac{5}{2},\frac{3}{2}\right):$



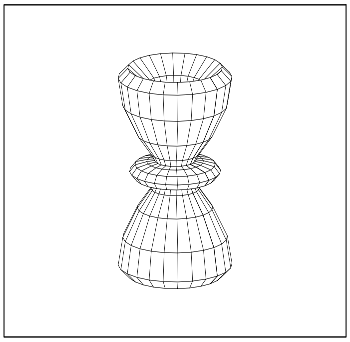
$d\left(\frac{5}{2},\frac{5}{2}\right), f\left(\frac{5}{2},\frac{5}{2}\right):$



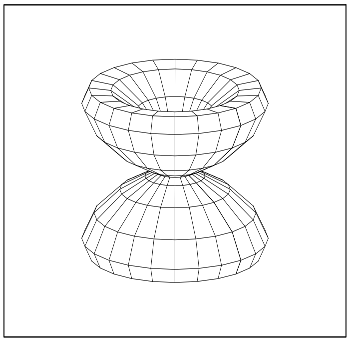
$d\left(\frac{7}{2},\frac{1}{2}\right), f\left(\frac{7}{2},\frac{1}{2}\right):$



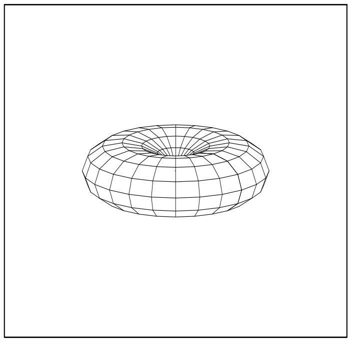
$d\left(\frac{7}{2},\frac{3}{2}\right), f\left(\frac{7}{2},\frac{3}{2}\right):$



$d\left(\frac{7}{2},\frac{5}{2}\right), f\left(\frac{7}{2},\frac{5}{2}\right):$



$d\left(\frac{7}{2},\frac{7}{2}\right), f\left(\frac{7}{2},\frac{7}{2}\right):$



Bibliography

[1] Hideki Yukawa. *Tabibito (The Traveler).* World Scientific, Singapore, 1982.

[2] J. G. Taylor. *Tributes to Paul Dirac.* Adam Hilger, Bristol, 1987.

[3] P. A. M. Dirac. Quantum mechanics of many-electron systems. *Proc. Roy. Soc. London*, A123:714, 1929.

[4] J. Almlöf and O. Gropen. Relativistic effects in chemistry. *Rev. Comp. Chem.*, page 8, 1996.

[5] P. Pyykkö. Relativistic effects in structural chemistry. *Chem. Rev.*, 88:563, 1988.

[6] S. S. Zumdahl. *Chemical Principles.* D.C.Heath and Company, Lexington,MA, 1995. pp. 548–9.

[7] P. Pyykkö. *Relativistic Theory of Atoms and Molecules. A Bibliography 1916–1985.* Lecture Notes in Chemistry 41. Springer, Berlin, 1986.

[8] P. Pyykkö. *Relativistic Theory of Atoms and Molecules II. A Bibliography 1986–1992.* Lecture Notes in Chemistry 60. Springer, Berlin, 1993.

[9] L. J. Norrby. Why is mercury liquid ? *J.Chem.Ed.*, 68:110–113, 1991.

[10] REHE newsletters are found at WWW-site http://pcgate.thch.uni-bonn.de/tc/hess/esf/nl.html.

[11] H. M. Quiney, I. P. Grant, and S. Wilson. On the relativistic many-body perturbation theory of atomic and molecular electronic structure. *Lecture Notes Chem.*, 52:307–344, 1989.

[12] I. P. Grant and H. M. Quiney. Foundations of the relativistic theory of atomic and molecular structure. *Adv. At. Mol. Phys.*, 23:37–85, 1988.

[13] K. G. Dyall. Polyatomic molecular dirac-hartree-fock calculations with gaussian basis sets: Theory, implementation and applications. In G. L. Malli, editor, *Relativistic and Electron Correlation Effects in Molecules and Solids*, New York and London, 1994. Plenum Press.

[14] W. Kutzelnigg. Perturbation theory of relativistic corrections: 2. analysis and classification of known and other possible methods. *Z.Phys.D*, 15:27, 1990.

[15] B. A. Hess. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Phys.Rev.A*, 33:3742, 1986.

[16] B. A. Hess, C. M. Marian, and S. D. Peyerimhoff. *Ab-initio Calculation of Spin-Orbit Effects in Molecules.* Advanced Series in Physical Chemistry. World Scientific, 1995.

[17] O. Gropen. The relativistic effective core potential method. In S. Wilson, editor, *Methods in Computational Chemistry*, volume II, page 109. Plenum Press, 1987.

[18] U. Häusermann, M. Dolg, H. Stoll, H. Preuss, and P. Schwerdtfeger. Accuracy of energy-adjusted, quasirelativistic, ab inition psudopotentials: all-electron and pseudopotential benchmark calculations for mercury , mercury hydride (hgh) and their cations. *J. Chem. Phys.*, 100:917–925, 1994.

[19] R. B. Ross, S. Gayen, and W. C. Ermler. Ab initio relativistic effective core potentials with spin-orbit operators. v. ce through lu. *J. Chem. Phys.*, 100:8145–8155, 1994.

[20] Z. Barandiaran and L. Seijo. Quasirelativistic ab initio model potential calculations on the group iv hydrides (xh_2, xh_4; x = si,ge,sn,pb) and oxides (xo; x = ge,sn,pb). *J. Chem. Phys.*, 93:4049, 1994.

[21] A. D. McLean and Y. S. Lee. Relativistic effects on r_e and d_e in agh and auh from all-electron dirac-hartree-fock calcualtions. *J. Chem. Phys.*, 76:735, 1982.

[22] L. Visscher, O. Visser, H. Aerts, H. Merenga, and W. C Nieuwpoort. Relativistic quantum chemistry: the molfdir program package. *Comput. Phys. Comm.*, 81:120, 1994.

[23] A. K. Mohanty and E. Clementi. Dirac-fock self-consistent field method for closed-shell molecules with kinetic balance and finite nuclear size. *Int. J Quant. Chem.*, 39:429–432, 1991.

[24] L. Pisani and E. Clementi. Relativistic dirac-fock calculations for closed-shell molecules. *J. Comput. Chem.*, 15:466, 1994.

[25] T. Saue. Utvikling av Dirac-Hartree-Fock-program for molekylære beregninger. Master's thesis, University of Oslo, 1991.

[26] R. E. Stanton and S. Havriliak. Kinetic balance: A partial solution to the problem of variational safety in dirac calculations. *J. Chem. Phys.*, 81:1910, 1984.

[27] K. G. Dyall, I. P. Grant, and S. Wilson. Matrix representation of operator products. *J. Phys. B*, 17:493, 1984.

[28] J. Sucher. Foundations of the relativistic theory of many-electron bound states. *Int. J. Quant. Chem.: Quant. Chem. Symp.*, 25:3, 1984.

[29] K. G. Dyall. *Chem. Phys. Lett.*, 224:186, 1994.

[30] J. K. Lærdahl, T. Saue, and K. Fægri. Direct relativistic mp2: Properties of ground state cuf, agf and auf. *Theor. Chem. Acc.*, 97:177–184, 1997.

[31] L. Visscher, T. Saue, W. C. Nieuwpoort, K. Fægri, and O. Gropen. The electronic structure of the pth molecule: Fully relativistic configuration interaction calculations of the ground and excited states. *J. Chem. Phys.*, 99:6704, 1993.

[32] L. Visscher, K. G. Dyall, and T. J. Lee. Kramers-restricted closed-shell ccsd theory. *Int. J. Quant. Chem.: Quant. Chem. Symp.*, 29:411–419, 1995.

[33] H.J.Aa Jensen, K. G. Dyall, T. Saue, and K. Fægri. Relativistic 4-component multiconfigurational self-consistent field theory for molecules: Formalism. *J. Chem. Phys.*, 104:4083, 1996.

[34] E. Engel, H. Müller, C. Speicher, and R. M. Dreizler. In E. K. U. Gross and R. M. Dreizler, editors, *Density Functional Theory*, NATO ASI Series. Series B: Physics Vol. 337, New York, 1995. Plenum.

[35] W. Kutzelnigg. Origin and meaning of the fermi contact interaction. *Theor. Chim. Acta*, 73:173, 1988.

[36] A. H. R. Buller. Punch Magazine 19 December 1923.

[37] James Gleick. *Genius — Richard Feynman and modern physics.* Little, Brown and Company, London, 1992.

[38] A. Einstein. Zur elektrodynamik bewegter körper. *Ann.Phys.*, 17:891–921, 1905.

[39] A. Pais. *Subtle is the Lord ... The Science and the Life of Albert Einstein.* Oxford University Press, Oxford, 1982.

[40] R. E. Moss. *Advanced Molecular Quantum Mechanics.* Chapmann and Hall, London, 1973.

[41] H. Goldstein. *Classical Mechanics.* Addison-Wesley, Reading, Massachusetts, 1980.

[42] C. G. Adler. Does mass really depend on velocity , dad ? *Am.J.Phys.*, 55:739–743, 1987.

[43] L. B. Okun. The concept of mass. *Phys.Today*, 6:31–36, 1989.

[44] W. R. Theis. A simple classical derivation of spin-orbit couplings. *Z.Phys. A*, 290:355, 1979.

[45] Max Jammer. *The Conceptual Development of Quantum Mechanics.* McGraw-Hill, New York, 1966.

[46] Y. Aharonov. Observability of the sign change of spinors under 2π rotaions. *Phys.Rev.*, 158:1237–1238, 1967.

[47] H. H. Bernstein. Spin precession during interferometry of fermions and the phase factor asociated with rotations through 2π radians. *Phys. Rev. Lett.*, 18:110, 1967.

[48] S. A. Werner. Gravitational, rotational and topological quantum phase shifts in neutron interferometry. *Class.Quantum.Grav.*, 11:A207–A226, 1994.

[49] F. Halzen and A. D. Martin. *Quarks & Leptons.* John Wiley, New York, 1984.

[50] S. Goudsmit and G. E. Uhlenbeck. *Naturw.*, 13:953, 1925.

[51] A. Pais. *Inward Bound.* Clarendon Press, Oxford, 1986.

[52] W. Pauli. Zur quantenmechanik des magnetischen elektrons. *Z.Phys.*, 43:601–623, 1927.

[53] J.-M. Lévy-Leblond. Nonrelativistic particles and wave equations. *Commun. Math. Phys.*, 6:286, 1967.

[54] J.-M. Lèvy-Leblond. Galilei group and galilean invariance. In E. M. Loebl, editor, *Group Theory and Its Applications, volume II*, page 221. Academic Press, New York, 1971.

[55] J.-M. Lèvy-Leblond. The pedagogical role and epistemological significance of group theory in quantum mechanics. *Nuov.Cim.*, 4:99, 1974.

[56] P. A. M. Dirac. *History of Twentieth Century Physics*, page 109. Academic Press, New York, 1977.

[57] W. Pauli and V. Weisskopf. *Helv.Phys.Act.*, 7:709, 1934.

[58] P. A. M. Dirac. The quantum theory of the electron. *Proc. Roy. Soc. London*, A117:714, 1928.

[59] P. A. M. Dirac. The quantum theory of the electron. part ii. *Proc. Roy. Soc. London*, A118:351, 1928.

[60] B. L. van der Waerden. *Group Theory and Quantum Mechanics.* Springer, Berlin, 1974. (Translation of the German Original Edition: Die Grundlehren der mathematischen Wissenschaften Band 37, Die Gruppentheoretische Methode in der Quantenmechanik. Publisher: Verlag von Julius Springer, Berlin 1932).

[61] B. N. Kursunoglu and E. P. Wigner, editors. *Reminiscences about a great physicist: Paul Adrien Maurice Dirac.* Cambridge University Press, 1987.

[62] E. P. Wigner. *Group theory and its application to the quantum mechanics of atomic spectra.* Academic Press, New York, 1959.

[63] S. A. Werner. Observation of the phase shift of a neutron due to precession in a magnetic field. *Phys. Rev. Lett.*, 35:1053–1055, 1975.

[64] H. Rauch, A. Zeilinger, G. Badurek, A. Wilfing, W. Bauspiess, and U. Bonse. Verification of coherent spinor rotation of fermions. *Phys.Lett.*, 54:425–427, 1975.

[65] A. G. Klein and G. I. Opat. Observation of 2π rotation by fresenl diffractin of neutrons. *Phys. Rev. Lett.*, 37:238–240, 1976.

[66] M. E. Stoll. Explicit demonstration of spinor character for a spin-1/2 nucleus via nmr interferometry. *Phys.Rev. A*, 16:1521–1524, 1977.

[67] H. Weyl. Elektron und gravitation.i. *Z.Phys.*, 56:330–352, 1929.

[68] C. S. Wu, E. Ambler, W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys.Rev.*, 1413:1413–1414, 1957.

[69] H. Bethe. Termaufspaltung in kristallen. *Ann.Phys.*, 3:133–208, 1929.

[70] G. Arfken. *Mathematical Methods for Physicists.* Academic Press, San Diego, 1985.

[71] S. L. Altmann and P. Herzig. Double goups and projective representations. iii. improper groups. *Mol.Phys.*, 45:585–604, 1982.

[72] S. L. Altmann. *Rotations, Quaternions, and Double Groups.* Clarendon Press, Oxford, 1986.

[73] E. Wigner. Über die operation der zeitumkehr in der quantenmechanik. *Nachrichten der Akad. der Wissensch. zu Göttingen,II*, pages 546–559, 1932.

[74] G. Aucar, H. J. Aa. Jensen, and J. Oddershede. Operator representations in kramers bases. *Chem. Phys. Lett.*, 232:47, 1995.

[75] G. M. Dixon. *Division Algebras.* Kluwer Academic, 1994.

[76] W. Pauli. Zur quantenmechanik des magnetischen elektrons. *Z.Phys.*, 43:601–623, 1927. See footnote 2 on page 607.

[77] C. P. Poole and H. A. Farach. Pauli-dirac matrix generators of clifford algebras. *Found.Phys.*, 12:719, 1982.

[78] L. E. Dickson. *Linear Algebras.* Cambridge University Press, London, 1914.

[79] C. Lanczos. Zur kovarianten formulierung der diracschen gleichung. *Z.Phys*, 57:474–483, 1929.

[80] C. Lanczos. Die erhaltungssätze in der feldmäßige darstellung der diracschen theorie. *Z.Phys*, 57:484, 1929.

[81] A. W. Conway. Quaternion treatment of the relativistic wave equation. *Proc. Roy. Soc. London*, A162:145–154, 1937.

[82] P. Rastall. Quaternions in relativity. *Rev.Mod.Phys.*, 36:820–832, 1964.

[83] K. Morita. Quaternionic formulation of dirac theory in special and general relativity. *Prog.Theor.Phys.*, 70:1648–1665, 1983.

[84] W. Greiner. *Relativistic Quantum Mechanics.* Springer-Verlag, Berlin, 1990.

[85] P. A. M. Dirac. A theory of electrons and protons. *Proc. Roy. Soc. London*, A126:360, 1929.

[86] C. D. Anderson. Energies of cosmic-ray particles. *Phys.Rev.*, 41:405–21, 1932.

[87] H. W. Kugel and D. E. Murnick. The lamb shift in hydrogenic ions. *Rep.Prog.Phys.*, 40:297–343, 1977.

[88] P. J. Mohr. Lamb shift in high-z atoms. In G. L. Malli, editor, *Relativistic Effects in Atoms, Molecules and Solids*, New York, 1983. Plenum Press.

[89] J. P. Briand, P. Chevalier, P. Indelicato, K. P. Ziock, and D. D. Dietrich. Observation and measurement of $n = 2 \rightarrow n = 1$ transitions of hydrogenlike and heliumlike uranium. *Phys. Rev. Lett.*, 65:2761–2764, 1990.

[90] T. Travis and T. Benfey. August wilhelm hofmann – a centennial tribute. *Education in Chemistry*, pages 69–72, 1992.

[91] M. Born and J. R. Oppenheimer. *Ann.Phys.*, 84:457, 1927.

[92] M. E. Rose. *Relativistic Electron Theory.* John Wiley, New York, 1961.

[93] Y. Ishikawa and H. M. Quiney. On the use of an extended nucleus in dirac-fock gaussian basis set calculations. *Int. J. Quant. Chem.: Quant. Chem. Symp.*, 21:523–532, 1987.

[94] O. Visser, P. J. C. Aerts, D. Hegarty, and W. C. Nieuwpoort. The use of gaussian nuclear charge distributions for the calculation of relativistic electronic wavefunctions using basis set expansions. *Chem. Phys. Lett.*, 134:34–38, 1987.

[95] W. R. Johnson and G. Soff. *At.Data Nucl. Data Tables*, 33:609, 1985.

[96] K. G. Dyall, K. Fægri, and P. Taylor. Polyatomic molecular dirac-hartree-fock calculations with gaussian basis sets. In I. P. Grant and B. Gyoffry, editors, *The Effects of Relativity in Atoms, Molecules and the Solid State*, page 167, New York, 1991. Plenum Press.

[97] J. Mehra, editor. *The Physicist's Conception of Nature.* D.Reidel Publishing Company, Dordrecht-Holland, 1973.

[98] W. Kutzelnigg. Perturbation theory of relativistic corrections: 1. the non-relativistic limit of the dirac equation and a direct perturbation expansion. *Z.Phys.D*, 11:15, 1989.

[99] K. G. Dyall. An exact separation of the spin-free and spin-dependent terms of the dirac-coulomb-breit hamiltonian. *J. Chem. Phys.*, 100:2118, 1994.

[100] W. Kutzelnigg. Basis set expansion of the dirac operator without variational collapse. *Int. J Quant. Chem.*, 25:107, 1984.

[101] L. L. Foldy and S. A. Wouthuysen. On the dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys.Rev.*, 78:29, 1950.

[102] R. E. Moss. The reduction of the dirac equation to a nonrelativistic form. *Am.J.Phys.*, 39:1169–1172, 1971.

[103] E. van Lenthe, E. J. Baerends, and J. G. Snijders. Relativistic regular two-component hamiltonians. *J. Chem. Phys.*, 99:4597, 1993.

[104] S. R. Langhoff and C. W. Kern. Modern Theoretical Chemistry. Plenum Press, New York, 1977.

[105] W. G. Richards, H. P. Trivedi, and D. L. Cooper. *Spin-orbit coupling in molecules.* Clarendon Press, Oxford, 1981.

[106] C Marian. *Berechnung von Matrixelementen des Spin-Bahn- und Spin-Spin-Kopplungsoperators mit MRD-CI-Wellenfunktionen.* PhD thesis, Rheinischen Friedrich-Wilhelms-Universität zu Bonn, 1981.

[107] E. Schrödinger. Über die kräftefreie bewegung in der relativistischen quantenmechanik. *Sitzungsber. phys. math. Kl.*, page 418, 1930.

[108] K. Huang. On the zitterbewegung of the dirac equation. *Am.J.Phys.*, 20:479–484, 1952.

[109] A. O. Barut. The importance of spin. In A. Trautman and G. Furlan, editors, *Spinors in Physics and Geometry*, Singapore, 1987. World Scientific.

[110] A. O. Barut and A. J. Bracken. *Zitterbewegung* and the internal geometry of the electron. *Phys.Rev.D*, 23:2454–2463, 1981.

[111] D. Hestenes. The zitterbewegung interpretation of quantum mechanics. *Foundations of Physics*, 20:1213–1231, 1990.

[112] K. G. Dyall and K. Fægri. Finite nucleus effects on realtivistic energy corrections. *Chem. Phys. Lett.*, 201:27, 1993.

[113] J. Almlöf, K. Fægri, and H. H. Grelland. A variational approach to relativistic effects in lcao calculations. *Chem. Phys. Lett.*, 114:53–57, 1985.

[114] J. Karwowski, J. Styszynski, and W. H. E. Schwarz. A numerical study of the validity of the breit-pauli approxaimation. *J.Phys. B*, 24:4877–4886, 1991.

[115] C. Chang, M. Pelissier, and P. Durand. Regular two-component pauli-like effective hamiltonians in dirac theory. *Phys.Scr.*, 34:394, 1986.

[116] S. M. Blinder. The hyperfine interaction hamiltonian. *J.Mol.Spectr.*, 5:17–23, 1960.

[117] K. G. Dyall, I. P. Grant, C. T. Johnson, F. A. Parpia, and E. P. Plummer. Grasp: A general-purpose relativistic atomic structure program. *Comput. Phys. Comm.*, 55:425–456, 1989.

[118] M. Douglas and N. M. Kroll. Quantum electrodynamical corrections to the fine structure of helium. *Ann.Phys.*, 82:89, 1974.

[119] B. A. Hess. personal communication.

[120] B. A. Hess. Applicability of the no-pair equation with free particle projection operators to atomic and molecular structure calculations. *Phys.Rev.A*, 32:756, 1985.

[121] R. Samzow, B. A. Hess, and G. Jansen. The two-electron terms of the no-pair hamiltonian. *J. Chem. Phys.*, pages 1–7, 1992.

[122] R. D. Mattuck. *A guide to Feynman diagrams in the many-body problem.* McGraw-Hill, New York, 1976.

[123] R. Crandall, R. Whitnell, and R. Bettega. Exactly soluble two-electron atomic model. *Am.J.Phys.*, 52:438–442, 1984.

[124] M. Planck. Über das gesetz der energieverteilung im normalspektrum. *Ann.Phys.*, 4:553–563, 1901.

[125] G. N. Lewis. The conservation of photons. *Nature*, 118:874, 1926.

[126] P. Indelicato. Effects of the breit interaction on the 1s binding energy of superheavy elements. *J. Phys. B: At. Mol. Opt. Phys.*, 19:1719–1730, 1986.

[127] O. Gorceix, P. Indelicato, and J. P. Desclaux. Multiconfiguration dirac-fock studies of two-electron ions: I. electron-electron interaction. *J. Phys. B: At. Mol. Opt. Phys.*, 20:639–649, 1987.

[128] I. Lindgren. Relativistic many-body calculations. *Phys.Scr.*, T34:36–46, 1991.

[129] E. Lindroth, A.-M. Mårtensson-Pendrill, A. Ynnerman, and P. Öster. Self-consistent treatment of the breit interaction, with application to the electric dipole moment in thallium. *J. Phys. B: At. Mol. Opt. Phys.*, 22:2447–2464, 1989.

[130] E. Lindroth. Furhter analysis of the complete breit interaction. *Phys.Rev.A*, 39:3794–3802, 1989.

[131] J. J. Sakurai. *Advanced Quantum Mechanics.* Addison-Wesley, Reading, Massachusetts, 1967.

[132] F. Mandl and G. Shaw. *Quantum Field Theory.* Wiley-Interscience, Chichester, 1984.

[133] J. Sucher. Foundations of the relativistic theory of many-electron atoms. *Phys.Rev*, 22:348, 1980.

[134] J. Sucher. Relativistic many-electron hamiltonians. *Phys.Scr.*, 36:271, 1987.

[135] J. Sucher. QED and relativity in Atomic Physics. In J. P. Briand, editor, *Atoms in Unusual Situations*, New York, 1986. Plenum Press.

[136] W. H. Furry. On bound states and scattering in positron theory. *Phys.Rev.*, 81:115–124, 1951.

[137] J. B. Mann and W. R. Johnson. On the breit interaction in multi-electron atoms. *Phys.Rev.A*, 4:41, 1971.

[138] J. Hata and I. P. Grant. Comments on relativistic two-body interactions in atoms. *J. Phys. B: At. Mol. Opt. Phys.*, 17:107, 1984.

[139] G. Breit. The effect of retardation on the interaction of two electrons. *Phys.Rev.*, 34:553–573, 1929.

[140] J. A. Gaunt. The triplets of helium. *Proc. Roy. Soc. London*, A122:513, 1929.

[141] J. Sucher. On the choice of the electron-electron potential in relativistic atomic physics. *J. Phys. B: At. Mol. Opt. Phys.*, page L585, 1988.

[142] Z. V. Chraplyvy. Reduction of relativistic two-particle wave equations to approximate forms. i. *Phys.Rev.*, 91:388–391, 1953.

[143] Z. V. Chraplyvy. Reduction of relativistic two-particle wave equations to approximate forms. ii. *Phys.Rev.*, 92:1310–1315, 1953.

[144] W. A. Barker and F. N. Glover. Reduction of relativistic two-particle wave equations to approximate forms. iii. *Phys.Rev.*, 99:317–324, 1955.

[145] H. A. Bethe and E. E. Salpeter. *Quantum Mechanics of one- and two-electron atoms.* Plenum/Rosetta, New York, 1977.

[146] Keh-Ning Huang. Orbit-orbit interaction. *J. Chem. Phys.*, 71:3830–3836, 1979.

[147] Y.-K. Kim. personal communication.

[148] O. Visser, L. Visscher, P. J. C. Aerts, and W. C. Nieuwpoort. Relativistic all-electron molecular hartree-fock-dirac- (breit) calculations on ch_4, sih_4, geh_4, snh_4 and pbh_4. *Theor. Chim. Acta*, 81:405, 1992.

[149] M. Twain. *Life on the Mississippi.* 1874.

[150] H. H. Grelland. Einstein relativistic correction to the nucleus-nucleus interaction in molecules in the born-oppenheimer approximation. *J. Phys. B: At. Mol. Opt. Phys.*, 13:L389–L390, 1980.

[151] S. T. Epstein. *The Variational Method in Quatnum Chemistry.* Academic Press, New York, 1974.

[152] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic Structure Theory.* John Wiley & Sons, Ltd, Chichester, 2000.

[153] L. Visscher. personal communication.

[154] K. Dyall, P. Taylor, K. Fægri, and H. Partridge. All-electron molecular dirac-hartree-fock calculations: The group iv tetrahydrides ch_4, sih_4, geh_4, snh_4, and pbh_4. *J. Chem. Phys.*, 95:2583, 1991.

[155] L. Pisani. Relativistic effects on sixth group hydrides. *J. Chem. Phys.*, 101:3079–3084, 1994.

[156] S. Huzinaga and C. Arnau. Virtual orbitals in hartree-fock theory.ii. *J. Chem. Phys.*, pages 1948–51, 1971.

[157] W. J. Hunt and W. A. Goddard III. Excited states of h_2o using improved virtual orbitals. *Chem. Phys. Lett.*, pages 414–418, 1969.

[158] J. Almlöf and P. Taylor. *J. Chem. Phys.*, 86:4070, 1987.

[159] P. Palmieri, R. Tarroni, and S. Rettrup. Hartree-fock operators to improve virtual orbitals and configuration interaction energies. *J. Chem. Phys.*, pages 5849–5856, 1994.

[160] P. O. Löwdin. On the nonorthogonality problem. *Adv.Quant.Chem.*, 5:185, 1970.

[161] G. E. Brown and D. G. Ravenhall. On the interaction of two electrons. *Proc. Roy. Soc. London*, A208:552, 1951.

[162] M. H. Mittleman. Theory of relativistic effects on atoms: Configuration-space hamiltonian. *Phys.Rev.A*, 24:1167, 1981.

[163] W. Kutzelnigg. The relativistic many body problem in molecular theory. *Phys.Scr.*, 36:416, 1987.

[164] J. D. Talman. Minimax principle for the dirac equation. *Phys. Rev. Lett.*, 57:1091, 1986.

[165] L. LaJohn and J. D. Talman. Minimax variational solution of the dirac equation in molecular geometries. *Chem. Phys. Lett.*, 189:383, 1992.

[166] J.-L. Heully, I. Lindgren, E. Lindroth, and A.-M. Mårtensson-Pendrill. Comment on relativistic wave equations and negative-energy states. *Phys.Rev. A*, 33:4426, 1986.

[167] N. N. Greenwood and A. Earnshaw. *Chemistry of the Elements.* Pergamon Press, Oxford, 1984.

[168] Y.-K. Kim. Relativistic self-consistent-field theory for closed-shell atoms. *Phys.Rev.*, 154:1739, 1967.

[169] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in FORTRAN.* Cambridge University Press, 1992.