An Introduction to Relativistic Quantum Mechanics

Patrick Sahrmann, University of Chicago

Abstract

The influence of special relativity on quantum chemical phenomena can be understood entirely from the Dirac equation. The Dirac equation is the fully relativistic wave equation for the electron, which holds amongst its predictions spin and its coupling to magnetic fields. The Dirac equation will be examined in the low-speed limit to return a Schrodinger equation with relativistic perturbations. These scalar and vector perturbations, which explain s/p orbital contraction and d/f orbital destabilization, will be derived. The effect of special relativity on chemical bonding will also be discussed through an example.

Introduction

Special relativity and quantum mechanics have been intimately connected since the origin of quantum theory. The Schrodinger equation presented a milestone in the new quantum theory, having accurately predicted the energy spectra of hydrogen. However, there remained peaks within the spectrum of hydrogen that were not explained by the Schrodinger equation, referred to as the "fine structure". It was understood quickly that these peaks corresponded to inherently relativistic effects. For example, in the frame of the electron, the proton moves around the stationary electron. This induces a corresponding magnetic field in addition to the Coulombic force. The spin of the electron is related to its intrinsic magnetic moment which couples to the magnetic field, inducing what is referred to as spin-orbit coupling. Furthermore, the Schrodinger equation incorporates a non-relativistic definition of kinetic energy. These two relativistic effects are referred to as "vector" and "scalar" corrections, respectively.[1]

A formal union of special relativity with quantum mechanics was not established until the discovery of the Dirac equation. The Dirac equation presents a fully-relativistic analogue to the Schrodinger equation. Indeed, in the non-relativistic limit, the Dirac equation reduces to the Schrodinger equation. The Dirac equation also predicts the excess angular momentum present in fermions, i.e., their spin. In this sense, the Dirac equation provides a justification for the fourth quantum number of atomic systems.[2] Beyond the inclusion of spin, it was regarded that relativistic corrections to quantum chemistry were unnecessary. It was not until the 1970s that quantum chemists seriously considered relativistic effects.

As one climbs up higher in atomic number in the periodic table, uncharacteristic chemical phenomena are observed which can be attributed to special relativity. A paragon of relativistic effects arising in chemical phenomena is the "gold maximum", in which the extreme contraction of the 6s orbital of gold accounts for the unusual color and non-reactivity which has made gold so highly desired. A general trend observed is that relativistic effects tend to contract the s and p orbitals due to high velocities, and spin-orbit coupling destabilizes the d and f orbitals via large angular momentum in open-shell systems. It is now understood that the need for relativity in quantum chemistry lies in the details. Such examples include the high ionization potentials of the outermost s-orbital electrons in post-transition metals. [3] Additionally, it was only recently discovered that the voltage of lead acid batteries, which involve a reaction between lead oxide and sulfuric acid, is generated predominantly via relativistic effects.[4] Ultimately, these results stem from the physical consequences of the Dirac equation.

Theory

Relativistic Quantum Mechanics

The Schrodinger equation can be inferred from the ansatz that the wavelength and frequency of the wavefuction associated with a particle is proportional to the particle's momentum and energy, respectively. Under these assumptions, momentum and energy take the following operator definitions:

$$\mathrm{p} = -i\hbar\nabla$$

$$E = i\hbar\frac{\partial}{\partial t}$$

Therefore, given the non-relativistic energy-momentum relation:

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

This implies the Schrodinger equation:

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{-\hbar^2}{2m}\nabla^2\Psi + V(x)\Psi$$

Clearly, the Schrodinger equation is non-relativistic, as the relativistic energy-momentum relation states:

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

Where c is the speed of light. Therefore, substituting the corresponding operator definitions:

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

Rearranging produces the Klein-Gordon equation:

$$-\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \nabla^2\Psi = \frac{m^2c^2}{\hbar^2}\Psi$$

The Klein-Gordon equation possesses equal order derivatives in space and time ,and clearly demonstrates Lorentz covariance when written as:

$$-\partial_\mu\partial^\mu\Psi = \frac{m^2c^2}{\hbar^2}\Psi$$

where the 'four-gradient' is defined as:

$$\partial_\mu = \left(\frac{1}{c}\partial_t, \partial_x, \partial_y, \partial_z\right)$$

and the Minkowski metric is defined as:

$$\eta^{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

While the Klein-Gordon equation at first glance satisfies the necessary prerequisites for a relativistic quantum theory, there is a crucial breakdown when applying the Born interpretation. Note that the probability density in non-relativistic quantum mechanics is given by:

$$\rho = \psi^*\psi$$

and as a result, is always positive definite. A necessary requirement is that the relativistic quantum theory must also satisfy this condition, i.e., that there are no negative probabilities. The associated Klein-Gordon density is:

$$\rho = i\left(\psi^* \frac{\partial \Psi}{\partial t} - \psi \frac{\partial \Psi^*}{\partial t}\right)$$

Note that since the Klein-Gordon equation is a second-order PDE, the first time derivative must be specified. This implies solutions to the Klein-Gordon equation in which there are negative probability densities. For this reason, physicists scrapped the Klein-Gordon equation as a suitable relativistic quantum theory, and sought alternatives. It is now understood that the Klein-Gordon equation, as well as all relativistic wave equations, do not describe the equations of motion of matter waves, but of the corresponding quantum fields from which the particles emerge. The conserved density is no longer associated with probability, but of conservation of charge. This reinterpretation would not have been possible without the physical insight provided by the Dirac equation.

The Dirac Equation

The derivation of the Dirac equation begins with the following ansatz: suppose that the d'Alembertian operator can be written as the following

$$-\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} + \nabla^2 \Psi = \left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)\left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)\psi$$

To make the space and time components more symmetric, introduce a new set of objects:

$$\gamma_\mu = (D, -iA, -iB, -iC)$$

Then:

$$-\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} + \nabla^2 \Psi = \left(i\gamma^\mu \partial_\mu\right)\left(i\gamma^\mu \partial_\mu\right)\psi$$

In order to avoid cross-terms, the collection of gammas must obey the following algebraic properties:

$$\{\gamma^\mu, \gamma^\nu\} = 0 \qquad \{\gamma^\mu, \gamma^\mu\} = 2$$

Or, in summary:

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$$

It is clear that only matrices, not ordinary numbers, will satisfy the given algebra. Prior to the development of relativistic quantum mechanics, it was well-known that the wavefunction possessed a 'spin' degree of freedom in addition to its spatial components. The spin observables S_x, S_y, and S_z are often expressed in terms of the Pauli matrices:

$$S_i = \frac{\hbar}{2}\sigma_i$$

The Pauli matrices exhibit the following anticommutation and commutation relations:

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \qquad [\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$$

This suggests three of the gamma matrices may be expressed in terms of the Pauli matrices. However, the gamma matrices commute, whereas the Pauli matrices do not. In an attempt to seek vanishing commutation relations, let the gamma matrices be 4x4 matrices, where:

$$\gamma^i = \begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix}$$

Then evaluating the anticommutation relation:

$$\{\gamma^i, \gamma^j\} = \begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix}\begin{bmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{bmatrix} + \begin{bmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{bmatrix}\begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix}$$

$$\{\gamma^i, \gamma^j\} = \begin{bmatrix} -\sigma^i\sigma^j & 0 \\ 0 & -\sigma^i\sigma^j \end{bmatrix} + \begin{bmatrix} -\sigma^j\sigma^i & 0 \\ 0 & -\sigma^j\sigma^i \end{bmatrix}$$

$$\{\gamma^i, \gamma^j\} = \begin{bmatrix} -\{\sigma^i, \sigma^j\} & 0 \\ 0 & -\{\sigma^i, \sigma^j\} \end{bmatrix} = 2\,\eta^{ij}$$

Thus, upon introducing 4x4 matrices, the commutation relations have been satisfied for three out of the four gamma matrices. All that remains is to find the γ^0 matrix. Employing a general definition in terms of 2x2 matrices:

$$\{\gamma^0, \gamma^j\} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{bmatrix} + \begin{bmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$\{\gamma^0, \gamma^j\} = \begin{bmatrix} -B\sigma^j & A\sigma^j \\ -D\sigma^j & C\sigma^j \end{bmatrix} + \begin{bmatrix} \sigma^j C & \sigma^j D \\ -\sigma^j A & -\sigma^j B \end{bmatrix}$$

Therefore, in general:

$$\{\gamma^0, \gamma^j\} = \begin{bmatrix} \sigma^j C - B\sigma^j & A\sigma^j + \sigma^j D \\ -\sigma^j A - D\sigma^j & C\sigma^j - \sigma^j B \end{bmatrix}$$

The simplest method to determine γ^0 is to note that the identity matrix commutes with all matrices. Therefore if γ^0 is diagonal, and A and D are related to the identity matrix, then all commutation relations are satisfied. The anticommutation relations then reduce to:

$$\{\gamma^0, \gamma^j\} = \begin{bmatrix} 0 & A + D \\ -(A+D) & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

This is satisfied when A and D are opposites. Therefore, let $A = I_2$ and $D = -I_2$. Then:

$$\gamma^0 = \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}$$

Now that a set of gamma matrices have been found which satisfy the afore-mentioned algebraic properties, the Klein-Gordon equation can be expressed as:

$$\left(i\gamma^\mu\partial_\mu\right)^2 \psi = \left(\frac{mc}{\hbar}\right)^2 \Psi$$

Or, written compactly:

$$\left(i\hbar\gamma^\mu\partial_\mu - mc\right)\psi = 0$$

This is the Dirac equation, it is the fully relativistic wave equation for a spin 1/2 particle.

The Emergence of Spin

Note that with the introduction of the 4x4 gamma matrices, the wavefunction can no longer be viewed as a scalar function, but rather as a bispinor, i.e., a 4-component vector. The spinor can be partitioned into two spinors (two-component vectors) ψ_+ and ψ_- such that:

$$\psi = \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix}$$

Then the Dirac equation may be rewritten as:

$$\left(\begin{bmatrix} E & -c\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ c\boldsymbol{\sigma}\cdot\boldsymbol{p} & -E \end{bmatrix} - \begin{bmatrix} mc^2 & 0 \\ 0 & mc^2 \end{bmatrix} \right) \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix} = 0$$

The ψ_- then constitute negative-energy solutions to the Dirac equation. Dirac proposed the concept of a sea of electrons all occupying the negative energy spectrum, however in modern quantum field theory these negative energy solutions are recognized as antimatter. The question still remains as to what the remaining two components are for the positive and negative energy components of the wavefunction. Recall that the wavefunction of a particle possesses an extra spin degree of freedom. To relate this to spin, working in the nonrelativistic limit, where:

$$E \approx mc^2$$

The Dirac equation produces two separate equations:

$$E\psi_+ - c\boldsymbol{\sigma}\cdot\boldsymbol{p}\psi_- - mc^2\psi_+ = 0$$

$$-E\psi_- + c\boldsymbol{\sigma}\cdot\boldsymbol{p}\psi_+ - mc^2\psi_- = 0$$

Rearranging the second:

$$E\psi_- + mc^2\psi_- = c\boldsymbol{\sigma}\cdot\boldsymbol{p}\psi_+$$

Solving:

$$\psi_- = \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E + mc^2}\psi_+ \approx \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{2mc}\psi_+$$

Substituting into the first equation produces:

$$E\psi_+ - \boldsymbol{\sigma}\cdot\boldsymbol{p}\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{2m}\psi_+ - mc^2\psi_+ = 0$$

Finally, utilizing the idempotency of the Pauli matrices:

$$E\psi_+ - \frac{p^2}{2m}\psi_+ - mc^2\psi_+ = 0$$

Or, upon rearranging:

$$\left(E - mc^2 \right)\psi_+ = \frac{p^2}{2m}\psi_+$$

The nonrelativistic limit returns the Schrodinger equation, but the wavefunction is still a two vector. Note that in the free particle Hamiltonian as derived above, the Schrodinger equation is identical for each component. In the presence of a magnetic field, this is not the case, as the Dirac equation becomes, via minimal coupling:

$$\left(\begin{bmatrix} E & -c\boldsymbol{\sigma}\cdot(\boldsymbol{p} - e\boldsymbol{A}) \\ c\boldsymbol{\sigma}\cdot(\boldsymbol{p} - e\boldsymbol{A}) & -E \end{bmatrix} - \begin{bmatrix} mc^2 & 0 \\ 0 & mc^2 \end{bmatrix} \right) \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix} = 0$$

and the nonrelativistic limit produces:

$$\left(E - mc^2\right)\psi_+ = \frac{\left(\boldsymbol{\sigma}\cdot(p - eA)\right)^2}{2m}\psi_+$$

Evidently, the particle's spin couples to the external magnetic field, and the "vector" relativistic effect emerges. This is now the Pauli equation, which introduced spin to the Schrodinger equation in an *ad hoc* fashion. The union of special relativity with quantum mechanics via the Dirac equation implies then antimatter, spin, and the Zeeman effect, and remains one of the greatest achievements of theoretical physics.

Applications

Perturbation Theory

Relativistic corrections to quantum systems will now be approached via perturbation theory. Working with the complete Dirac equation for the electron:

$$E\psi_+ - \boldsymbol{\sigma}\cdot p\frac{c^2\boldsymbol{\sigma}\cdot p}{E + mc^2}\psi_+ - mc^2\psi_+ = 0$$

Then applying the nonrelativistic limit:

$$E\psi_+ - \frac{p^2}{(1+\gamma)m}\psi_+ - mc^2\psi_+ = 0$$

Rearranging and applying a Taylor series expansion:

$$E\psi_+ - mc^2\psi_+ = \frac{p^2}{(1+\gamma)m}\psi_+ \approx \frac{p^2}{2m}\psi_+ - \frac{1}{8}\left(\frac{p^4}{m^3c^2}\right)\psi_+$$

Now consider the Hydrogen atom:

$$E\psi_+ - mc^2\psi_+ = \frac{p^2}{2m}\psi_+ - \frac{1}{8}\left(\frac{p^4}{m^3c^2}\right)\psi_+ - \frac{1}{4\pi\varepsilon_0 r}\psi_+$$

The mc^2 term is just a constant. Therefore by shifting the energy by mc^2 term produces the Schrodinger equation for the hydrogen atom with an additional perturbation.

$$H_{Hyd}\psi - \frac{1}{8}\left(\frac{p^4}{m^3c^2}\right)\psi_+ = E\psi$$

Applying perturbation theory:

$$E_{rel} = \int \Psi^* H_{rel}\psi\, dV = -\frac{1}{8m^3c^2}\langle p^4\rangle$$

By re-expressing the perturbation as:

$$\langle p^4\rangle = \langle p^2 p^2\rangle = 2m\langle (E - V)^2\rangle$$

The answer can be expressed via the virial theorem:

$$E_{rel} = -\frac{E_n^2}{2mc^2}\left(\frac{4n}{l + \frac{1}{2}} - 3\right)$$

where E_n are the energy states of the hydrogen atom:

$$E_n = -\frac{m}{2\hbar^2} \left(\frac{Ze^2}{4\pi\varepsilon_0} \right)^2 \frac{1}{n^2}$$

Therefore relativistic effects due solely to the speed of the electron scale roughly with the square of the nuclear charge. Thus it is expected that relativistic effects become more relevant traversing down the periodic table. The "scalar" relativistic effect emerges then from the higher nuclear charge contracting inner shell orbitals.

Copper: An Example

Relativistic effects have a tendency to shrink equilibrium bond lengths. Take as an example, the Cu-Cu bond. The Cu-Cu molecule has an experimentally observed bond length of 222 picometers, which disagrees significantly with that calculated via non-relativistic HF.[5] The non-relativistic equilibrium bond length can be calculated by varying the spatial distance between two Copper atoms and calculating the energy. Using a set of bond distances from the roots of the sixth-order Chebyshev polynomial:

> $with(QuantumChemistry);$

$[AOLabels, ActiveSpaceCI, ActiveSpaceSCF, AtomicData, BondAngles, BondDistances,$ **(4.2.1)**
$\quad Charges, ChargesPlot, CorrelationEnergy, CoupledCluster, DensityFunctional,$
$\quad DensityPlot3D, Dipole, DipolePlot, Energy, FullCI, GeometryOptimization, HartreeFock,$
$\quad Interactive, Isotopes, MOCoefficients, MODiagram, MOEnergies, MOIntegrals,$
$\quad MOOccupations, MOOccupationsPlot, MOSymmetries, MP2, MolecularData,$
$\quad MolecularGeometry, NuclearEnergy, NuclearGradient, Parametric2RDM, PlotMolecule,$
$\quad Populations, RDM1, RDM2, ReadXYZ, SaveXYZ, SearchBasisSets, SearchFunctionals,$
$\quad SkeletalStructure, Thermodynamics, Variational2RDM, VibrationalModeAnimation,$
$\quad VibrationalModes, Video]$

> $bond\_distances := map(x \rightarrow x/10 + 2.35, [fsolve(expand(ChebyshevT(6, x)))]);$
$bond\_distances := [2.25340742, 2.27928932, 2.32411810, 2.37588190, 2.42071068,$ **(4.2.2)**
$\quad 2.44659258]$

Varying the bond distance around roughly around the equilibrium bond length

> $molecules := [seq([["Cu", 0, 0, 0], ["Cu", 0, 0, R]], R \textbf{ in } bond\_distances)];$
$molecules := [[["Cu", 0, 0, 0], ["Cu", 0, 0, 2.25340742]], [["Cu", 0, 0, 0], ["Cu", 0, 0,$ **(4.2.3)**
$\quad 2.27928932]], [["Cu", 0, 0, 0], ["Cu", 0, 0, 2.32411810]], [["Cu", 0, 0, 0], ["Cu", 0, 0,$
$\quad 2.37588190]], [["Cu", 0, 0, 0], ["Cu", 0, 0, 2.42071068]], [["Cu", 0, 0, 0], ["Cu", 0, 0,$
$\quad 2.44659258]]]$

Calculating the energies for each geometry:

> $energies := [seq(Energy(molecule, method = HartreeFock, basis = "cc\text{-}pVDZ"), molecule$
$\quad \textbf{in } molecules)];$
$energies := [-3277.93836385, -3277.93934052, -3277.94061085, -3277.94150580,$ **(4.2.4)**
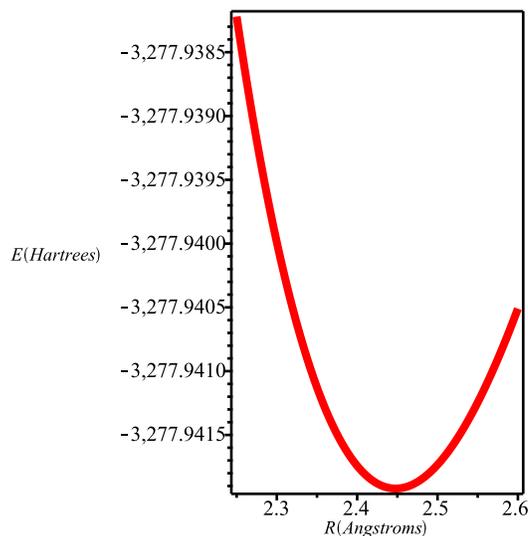$\quad -3277.94186439, -3277.94191997]$

Using polynomial interpolation to generate a polynomial in terms of the bond distance R

> $pes := interp(bond\_distances, energies, R);$
$pes := -0.10406127 R^5 + 1.34755941 R^4 - 7.04841872 R^3 + 18.64948298 R^2 - 24.98061971 R$ **(4.2.5)**

$$-3264.39471599$$

Plotting the potential energy curve

> $plot(pes(R), R = 2.25 ..2.6, axes = boxed, labels = ['R(Angstroms)', 'E(Hartrees)'], color = red,$
$thickness = 3);$

The minimum will occur where the derivative is zero with respect to the bond distance, therefore:

> $eq := diff(pes, R) = 0;$

$$eq := -0.52030637\,R^4 + 5.39023766\,R^3 - 21.14525615\,R^2 + 37.29896596\,R - 24.98061971 \qquad \textbf{(4.2.6)}$$
$$= 0$$

Solving the resulting equation yields the equilibrium bond length

> $R\_eq := fsolve(eq, R = 2.3 ..2.5);$

$$R\_eq := 2.44784626 \qquad \textbf{(4.2.7)}$$

Therefore the equilibrium bond length according to HF calculations with a cc-pVDZ basis set is 245 picometers. In a study of the relativistic contribution to the bond length of Cu-Cu, Hartree-Fock theory using a primitive Gaussian basis set incorporating 14s9p5d basis elements, it was predicted that the Cu-Cu bond length would have a minimum at approximately 242 picometers. While the overestimated bond length can be attributed to inadequate evaluation of electron correlation, incorporating relativistic effects with the same method and basis set significantly decreased the bond length. Specifically, via the introduction of relativistic terms to the Hamiltonian, the Cu-Cu bond length was shortened to 237 picometers.

Discussion

Modern Research

Currently, there exist multiple approaches to incorporating relativistic effects in quantum chemistry programs. One example is to introduce a new Hamiltonian apart from the Schrodinger Hamiltonian, referred to as the two-component approach. Similarly, single and multi-configurational Dirac-Fock methods exist, which take into account the entire four-component spinor. Lastly, pseudopotentials are widely used to incorporate relativistic effects.

Conclusion

It is evident that to a certain extent, chemical phenomena and special relativity are intertwined. Spin, which has deeply influenced electronic structure since the origins of quantum chemistry (e.g., Aufbau and Pauli Exclusion), must be regarded as an entirely relativistic effect. Beyond this, the effects of relativity on quantum chemistry are diverse and subtle. A particular challenge of quantum chemistry is that the realm of non-marginal relativistic effects overlaps significantly with that of strongly correlated systems. Relativistic quantum chemistry remains an active field, which continues to discover the influence of special relativity on chemical species composed of elements ranging from the 3d block to hypothetical elements not yet discovered.

References

1. Pyykko, P. Relativistic Effects in Structural Chemistry. *Chemical Reviews* **1988**, *88* (3), 563–594. https://pubs.acs.org/doi/abs/10.1021/cr00085a006

2. Dirac, P.A.M. The Quantum Theory of the Electron. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1928**, *117* (778), 610–624. https://royalsocietypublishing.org/doi/10.1098/rspa.1928.0023

3. C. K. Jørgensen. *Modern Aspects of Ligand Field Theory.* North-Holland Publishing Co. **1971**.

4. Ahuja, R.; Blomqvist, A.; Larsson, P.; Pyykkö, P.; Zaleski-Ejgierd, P. Relativity and the Lead-Acid Battery. *Physical Review Letters* **2011**, *106* (1). https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.106.018301

5. Martin, R. L. On Relativistic Contributions to the Bonding in Cu_2. *The Journal of Chemical Physics* **1983**, *78* (9), 5840–5842. https://aip.scitation.org/doi/10.1063/1.445430